The Association Between Early-life Written Language Skills and Late-life Cognitive Resilience to Alzheimer's Disease

by

Danielle Olivia Fearon

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Science

in

Health Studies and Gerontology

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

As the population ages, projections suggest that the number of individuals living with age-related diseases such as Alzheimer's disease will increase. Prevention of Alzheimer's disease is a major priority since there is currently no cure for the disease. Cognitive resilience is a hypothetical construct designed to explain why some individuals manage to avoid cognitive changes despite the presence of Alzheimer neuropathology. Educational attainment is one of the well-documented examples of building cognitive resilience since high levels of educational attainment have been associated with delayed onset of cognitive impairment. Written language skills developed in early life may reflect the development of early intellect and are essential to educational attainment. Weak early-life written language skills (i.e., low idea density and low grammatical complexity) have been associated with poor cognitive function in later life. However, there is limited understanding of the influence of written language skills and their potential contribution to cognitive resilience.

This research aimed to assess the association between written language skills and cognitive resilience using data from the Nun Study. The Nun Study is a longitudinal study of aging in religious sisters who were a minimum of 75 years of age at baseline. Idea density and grammatical complexity were determined using coded autobiographies. Autobiographies were obtained from archival records and were written at a mean age of 22 years. Cognitive resilience was operationalized based on whether individuals met the clinical diagnosis of dementia at last assessment prior to death according to DSM-IV criteria while fulfilling Consortium to Establish a Registry for Alzheimer's Disease (CERAD) neuropathologic criteria ("definite" or "probable") or National Institute on Aging and Reagan Institute (NIA-RI) neuropathologic criteria ("definite", "intermediate" or "high" likelihood) for Alzheimer's disease. Analyses included descriptive analyses (univariate and bivariate) as well as logistic regression models. The purpose of this project was to strengthen current knowledge on the potential association between early-life written language skills and late-life resilience to cognitive impairment. This study also aimed to better understand the implications of indicators of cognitive and brain reserve on this potential relationship.

Based on descriptive and multivariable analyses, a relationship between written

language skills (idea density and grammatical complexity) was found particularly in the CERAD sample where cognitive resilience was defined using CERAD criteria for Alzheimer neuropathology. In logistic regression models adjusting for standard covariates (age and *APOE*), low idea density was associated with decreased likelihood of cognitive resilience (Odds Ratio (OR): 0.15, 95% Confidence Interval (CI): 0.02-0.72). These findings meant that higher idea density (vs. low) was associated with six times greater odds of cognitive resilience. Similarly, low grammatical complexity was significantly associated with cognitive resilience in adjusted models for age and *APOE* (OR: 0.13, 95% CI: 0.03-0.50). That is, the odds of cognitive resilience in later life increased seven-fold among those with higher grammatical complexity compared to those with low grammatical complexity. Further analyses also suggested that grammatical complexity remained a significant predictor of cognitive resilience in the presence of indicators of cognitive (education) and brain (cerebral infarcts and cortical atrophy) reserve. In comparison, idea density was significant when separately adjusted for presence or number of infarcts along with standard covariates. However, idea density was not significant in a few full models (e.g., including adjustments for standard covariates (age and *APOE*), cortical atrophy and presence of infarcts, or standard covariates and education). These findings suggested the strong influences of both education and structural brain changes on the relationship between idea density and cognitive resilience. Future studies should aim to assess whether other forms of writing from early life (e.g., written language in social media) can also be associated with cognition in later life. Findings from this research contribute to the understanding of cognitive resilience and provide the foundation for further exploration into the influence of written language on the prevention of Alzheimer's disease.

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Suzanne Tyas. I appreciate your support, guidance and commitment. Without you, I would not have been able to select a thesis topic or develop research skills necessary for the future. Thank you for helping me discover my interests in epidemiology and the application of statistics to health-related issues across the life course. I also thank my committee members, Dr. Colleen Maxwell and Dr. Randy Harris. I appreciate how committed you both were in providing helpful feedback, support, and advice throughout this process. I also thank you both for your time, and commitment.

I thank my fellow graduate students I met throughout this process. I am grateful to have met such wonderful people in my program, in our research group as well as at the graduate house. I appreciate your support and friendships throughout my graduate studies.

I would also like to thank my parents and my two brothers. You have all taught me the importance of reaching your goals and I am so thankful to have such a loving, caring and supportive family. I also thank my grandmother, Olive who always provides me with support, valuable advice and endless prayers.

Last but in no way least, I thank my husband, Blake Madill. You have been my rock and my biggest fan throughout this degree. I thank you for providing me with unconditional love and support, even when I felt like giving up. I am beyond lucky to have such an amazing life partner and your passion for mathematics inspires me each and every day. Thank you for everything you have given me, I am forever grateful.

Table of Contents

AUTHOR'S DECLARATION ii

Abstract iii

Acknowledgements v

Table of Contents vi

List of Figures ix

List of Tables x

List of Abbreviations xii

Chapter 1: Introduction 1

Chapter 2: Literature Review 4

2.1 Alzheimer's Disease 4

2.1.1 Public Health Impact 4

2.1.2 Etiology 5

2.1.3 Diagnosis 6

2.1.4 Risk Factors 8

2.2 Cognitive Resilience 12

2.2.1 Brain and Cognitive Reserve 12

2.2.2 Resilience 13

2.3 Language Skills 14

2.3.1 Childhood Development of Language Skills 14

2.3.2 Multilingualism 16

2.3.3 Written Language Skills 16

2.3.3.1 Oral Versus Written Language Skills 16

2.3.3.2 Idea Density and Grammatical Complexity 19
2.3.3.3 Written Language Skills and Cognition 20
2.3.4 Summary of Written Language Skills 21
Chapter 3: Study Rationale and Research Questions 23
3.1 Study Rationale 23
3.2 Research Questions 24
Chapter 4: Methodology 26
4.1 Literature Search Strategy 26
4.2 Data Source: The Nun Study 27
4.2.1 Background 27
4.2.2 Participants 27
4.2.3 Autobiographies 28
4.3 Analytic Sample 28
4.4 Measures 31
4.4.1 Exposure Measures 31
4.4.2 Outcome Measures 33
4.4.3 Covariates 33
4.5 Analysis 33
4.5.1 Descriptive Analyses 33
4.5.2 Multivariable Analyses 34
4.6 Ethics 34
Chapter 5: Results 36
5.1 Descriptive Analyses 36

5.2 Multivariable Analyses 39
5.2.1 Research Question 1 39
5.2.2 Research Question 2 44
5.2.3 Research Question 3 49
5.2.4 Research Question 4 52
Chapter 6: Discussion 74
6.1 Summary of Findings 75
6.2 Limitations 80
6.3 Strengths 82
6.4 Implications and Future Directions 82
References 85
Appendix A: Literature Search Strategies 94
Appendix B: Literature Summary Table 95
Appendix C: Description of Grammatical Complexity and Idea Density 103
Appendix D: Sensitivity Analyses Using a Sample of Participants that Met Either CERAD or NIA-RI Criteria 105
Appendix E: Assessment of Selection Bias in Analytic Samples 120

List of Figures

Figure 1. Flowchart of Analytic Sample 30

Figure 2. Timeline of Relevant Nun Study Measures 32

List of Tables

Table 1a. Participant Characteristics by Cognitive Resilience Status, CERAD Criteria 37

Table 1b. Participant Characteristics by Cognitive Resilience Status, NIA-RI Criteria 38

Table 2a. Participant Characteristics Stratified by Idea Density, CERAD Criteria 40

Table 2b. Participant Characteristics Stratified by Grammatical Complexity, CERAD Criteria 41

Table 2c. Participant Characteristics Stratified by Idea Density, NIA-RI Criteria 42

Table 2d. Participant Characteristics Stratified by Grammatical Complexity, NIA-RI Criteria 43

Table 3a. The Effect of Age and Apolipoprotein E on the Association of Idea Density and Cognitive Resilience, CERAD Criteria 46

Table 3b. The Effect of Age and Apolipoprotein E on the Association of Grammatical Complexity and Cognitive Resilience, CERAD Criteria 47

Table 3c. The Effect of Age and Apolipoprotein E on the Association of Grammatical Complexity and Cognitive Resilience, NIA-RI Criteria 48

Table 4a. The Impact of Education on the Association of Idea Density and Cognitive Resilience using Firth Logistic Regression, CERAD Criteria 50

Table 4b. The Impact of Education on the Association of Grammatical Complexity and Cognitive Resilience using Firth Logistic Regression, CERAD Criteria 51

Table 5a. The Association Between Idea Density and Cognitive Resilience Adjusted for Presence of Cerebral Infarcts, CERAD Criteria 54

Table 5b. The Association Between Grammatical Complexity and Cognitive Resilience Adjusted for Presence of Cerebral Infarcts, CERAD Criteria 55

Table 5c. The Association Between Idea Density and Cognitive Resilience Adjusted for Number of Cerebral Infarcts, CERAD Criteria 56

Table 5d. The Association Between Grammatical Complexity and Cognitive Resilience Adjusted for Number of Cerebral Infarcts, CERAD Criteria 57

Table 6a. The Association Between Idea Density and Cognitive Resilience Adjusted for Cortical Atrophy, CERAD Criteria 59

Table 6b. The Association Between Grammatical Complexity and Cognitive Resilience Adjusted for Cortical Atrophy, CERAD Criteria 60

Table 7a. The Association Between Idea Density and Cognitive Resilience Adjusted for Presence of Cerebral Infarcts and Cortical Atrophy, CERAD Criteria 63

Table 7b. The Association Between Grammatical Complexity and Cognitive Resilience Adjusted for Presence of Cerebral Infarcts and Cortical Atrophy, CERAD Criteria 64

Table 7c. The Association Between Idea Density and Cognitive Resilience Adjusted for Number of Cerebral Infarcts and Cortical Atrophy, CERAD Criteria 65

Table 7d. The Association Between Grammatical Complexity and Cognitive Resilience Adjusted for Number of Cerebral Infarcts and Cortical Atrophy, CERAD Criteria 66

Table 8. Summary of Findings on the Association Between Early-life Written Language Skills and Cognitive Resilience 68

List of Abbreviations

AD	Alzheimer's disease
APOE	Apolipoprotein E
CERAD	Consortium to Establish a Registry for Alzheimer's Disease
CI	Confidence intervals
CIND	Cognitive Impairment, Not Dementia
DSM	Diagnostic and Statistical Manual
MCI	Mild cognitive impairment
MMSE	Mini-Mental State Examination
MRI	Magnetic resonance imaging
NFT	Neurofibrillary tangle
NIA-RI	National Institute on Aging and Reagan Institute
NINCDS-ADRDA	National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association
NP	Neuritic plaque
OR	Odds ratio
SD	Standard deviation

Chapter 1
Introduction

Population aging presents many societal and economic challenges. It is projected that the global percentage of those over the age of 60 will increase from 12% to 22% between the years of 2015 and 2050 (World Health Organization (WHO), 2015). Projections of population aging in Canada suggest that the proportion of older adults (individuals ages 65 and older) will grow from 14% (4.8 million) to 25% (10.4 million people) by the year 2036 (Canadian Institute for Health Information (CIHI), 2011). Population aging has major economic impacts on society that are reflected by per capita spending on health care services. In 2009, the per capita annual Canadian expenditure for individuals 65 and older was $11,196, which was 4.5 times greater than the per capita spending on individuals between the ages of 20 to 64 ($2,494) (CIHI, 2011). Since population aging has societal and economic impacts, addressing underlying age-related conditions is a major priority.

Of the age-related diseases, dementias are common and can be devastating to both individuals and their care partners. Dementia is an umbrella term for debilitating progressive cognitive disorders that include symptoms such as cognitive decline and memory loss (Prince et al., 2013). The most common form of dementia is Alzheimer's disease (AD) (Tyas & Gutmanis, 2015). In 1906, AD was described by Dr. Alois Alzheimer who identified the distinctive accumulation of neuropathological plaques and tangles (Hardy & Higgins, 1992). AD is associated with changes in cognitive and functional abilities, emotions and behaviour, and results in premature death (Prince et al., 2013).

Strategies that build resilience to cognitive impairments and reduce the impact of AD are of increasing interest. Cognitive resilience describes the cognitive and structural mechanisms that allow some individuals to maintain cognitive function in the presence of Alzheimer neuropathology (Stern, 2002). Some early-life factors have been associated with increasing cognitive resilience in older life.

The association between low educational attainment and an increased risk of developing AD and dementia in later life is one of the most documented examples of cognitive resilience (Stern et al., 1994). Theories of cognitive resilience also suggest that individuals with higher levels of educational attainment should have delayed onset of

cognitive impairment.

Linguistic ability may be a stronger indicator of early-life cognitive ability than education (Snowdon et al., 1996). Written language skills can be understood as a proxy measure of an individual's ability to process, encode and retrieve information. Early-life factors may also reflect the development of early intellect and conceptual ideas (Snowdon et al., 2000). For analysis of written language skills (idea density and grammatical complexity), autobiographies from the Nun Study (a longitudinal study of aging among religious sisters) were used. Autobiographies were completed while participants were a mean age of 22 (Snowdon et al., 2000). In the context of this thesis on late-life cognitive impairment, the term "early life" referred to young adulthood. These writing skills developed in early life are essential to educational attainment and success throughout life (Snowdon et al., 1996).

Existing literature suggests an association of low levels of early-life written language skills with poor overall cognitive function and an increased risk of dementia in later life (Snowdon et al., 2000). However, written language skills have not been analyzed with respect to cognitive resilience.

The overall purpose of the present study was to enhance current knowledge on the association between early-life written language skills and later-life cognitive resilience. The first aim of this study was to describe how cognitive resilience varied by level of written language skills. The second aim was to assess the association between written language skills and cognitive resilience. The third aim was to determine if the association between written language skills and cognitive resilience held in the presence of indicators of cognitive reserve, and the fourth aim was to assess whether the association held in the presence of indicators of brain reserve.

To test these research questions, analyses were conducted using secondary data from the Nun Study. The Nun Study is a longitudinal study of aging in members of a religious congregation in the United States (Snowdon et al., 1997). Measures of written language skills (idea density and grammatical complexity) were retrieved from autobiographies in archival records (Danner et al., 2001). Cognitive resilience was operationalized based on whether individuals met the clinical diagnosis of dementia at last assessment prior to death according

to DSM-IV criteria while fulfilling CERAD neuropathologic criteria ("definite" or "probable") or National Institute on Aging and Reagan Institute (NIA-RI) neuropathologic criteria ("definite", "intermediate" or "high" likelihood) for Alzheimer's disease. The present research also accounted for relevant covariates and confounding variables, such as age at death, education and apolipoprotein E (*APOE*).

This research will have many practical implications. Firstly, cognitive resilience is a complex outcome. Thus, strengthening our understanding of factors that develop resilience is beneficial to reduce the impact of age-related diseases such as AD. Additionally, early life is an important time frame in terms of development, and findings from this research could support the need for continued focus on early language skills and beneficial early-life circumstances. Written language skills, is a modifiable risk factor that could be strengthened to develop resilience in later life. The present research investigated whether early-life written language skills predict cognitive resilience in later life. A better understanding of the potential association between written language skills and cognitive resilience could inform public health interventions, such as implementing strategies to improve the development of early intellect, language skills (e.g., reading and writing) and conceptual processing to reduce the burden of AD in late life.

Chapter 2
Literature Review

2.1 Alzheimer's Disease

2.1.1 Public Health Impact

Age-related diseases such as dementias have major public health impacts (Tyas & Gutmanis, 2015). In Canada, it is estimated that 25,000 individuals are diagnosed with a form of dementia annually. However, many cases of dementia go undiagnosed, presenting further challenges such as lack of appropriate interventions (Tyas & Gutmanis, 2015). It is estimated that there are currently 564,000 Canadians currently living with dementia (Alzheimer Society of Canada, 2016). The number of Canadians living with dementia is expected to rise to 937,000 (a 66% increase) by the year 2031 (Alzheimer Society of Canada, 2016). Of these newly diagnosed individuals, women account for approximately 65 percent (Alzheimer Society of Canada, 2017).

Dementia is a general term used to describe a category of progressive cognitive disorders. Dementias are associated with a variety of symptoms such as declines in memory and a reduced ability to perform activities of daily living (Alzheimer Society of Canada, 2017). The term "dementia" itself is not a specific disease, but reflects a clinical picture caused by a variety of conditions such as AD, vascular dementia or Parkinson's disease (Alzheimer Society of Canada, 2017).

Changes to one's physical and mental abilities can occur as a result of dementia, which can reduce independence and quality of life. These changes in lifestyle and associated challenges can be difficult for both the individual and their care partners (WHO, 2017). In addition to lifestyle changes, financial burdens also pose challenges. In Canada, the financial cost of dementia is approximately $570 billion, further emphasizing the tremendous burden of dementia on individuals, care partners, and society in general (Alzheimer Society of Canada, 2010).

Dementias can be either reversible (i.e., curable with treatment) or irreversible (i.e., not currently curable with treatment). Reversible dementias are generally secondary symptoms and the result of overarching issues such as infections, depression, or the use of

certain medications (Tyas & Gutmanis, 2015). Dementias that are reversible can generally be addressed by using appropriate medical interventions. However, prevention of factors that contribute to reversible dementias remain important. With respect to irreversible dementias, further understanding of the underlying neuropathologic changes and prominent features of the disease are necessary in preventing their occurrence.

2.1.2 Etiology

Of the irreversible dementias, AD is the most common form (Tyas & Gutmanis, 2015; Alzheimer Society of Canada, 2017). AD is a progressive disorder affecting the brain and causing changes in memory, reasoning, problem solving skills, and ability to perform activities of daily living (National Institute on Aging, 2016).

AD was described by Dr. Alois Alzheimer, who first identified the distinctive accumulation of neuropathological lesions: neuritic plaque (NP) deposits and neurofibrillary tangles (NFT) (Carrillo, Thies, & Bain, 2012; SantaCruz et al., 2011). NPs consist of insoluble deposits of toxic Aβ amyloid protein peptide and are located between clusters of nerve cells (Hyman et al., 2012; National Institute on Aging, 2016). Neurofibrillary tangles are the intracellular collection of abnormally twisted tau protein (Hyman et al., 2012; National Institute on Aging, 2016). Phosphate molecules are naturally attached to tau protein and are responsible for stabilizing and binding to microtubules. In cases where additional phosphate molecules attach to tau, hyperphosphorylation occurs (Hyman et al., 2012; National Institute on Aging, 2016). Hyperphosphorylation results in disruptions in transport and communication between cells and can cause cell death and brain atrophy (National Institute on Aging, 2016).

Dr. Alois Alzheimer concluded that these forms of lesions (NPs and NFTs) were the root cause of AD (Alzheimer's Association, 2017). However, there are currently a variety of major hypotheses on the origin of Alzheimer neuropathology. The amyloid cascade hypothesis is one of these major hypotheses and suggests a deposition of amyloid β protein (a main component in NPs) is the cause of AD neuropathology (Hardy & Higgins, 1992). Other hypotheses suggest mechanisms such as oxidative stress, calcium dysregulation and genetics as potential causes of AD neuropathology (Tyas & Gutmanis, 2015). Regardless of the

hypothesis, NPs and NFTs remain the most prominent pathologic features of AD (National Institute on Aging, 2016).

2.1.3 Diagnosis

Clinical Criteria

For a gold standard diagnosis of AD, both the presence of dementia during life (clinical criteria) and the verification of post-mortem AD neuropathology (neuropathologic criteria) are required. A clinical diagnosis of AD uses standard tests and clinical assessments of memory and attention (National Institute on Aging, 2016). Urine or blood samples can also be taken to rule out other potential causes of illness. Other methods, such as brain scans or magnetic resonance imaging (MRI), are used to inform diagnosis of AD in living individuals (National Institute on Aging, 2016).

The Diagnostic and Statistical Manual of Mental Disorders versions IV and V (DSM-IV/V) also provide direction on the diagnosis of dementia (Alzheimer's Association, 2016). Some of the criteria include memory impairment, disturbance in language use (aphasia), motor function impairment (apraxia), inability to recognize objects (agnosia), and changes in executive function skills (Alzheimer's Association, 2016). Other clinical criteria from the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) and the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) provide guidelines on identifying AD during life. All of these clinical diagnoses are dependent on a clinical interview assessing AD symptoms, level of cognition, and physical and neurological exams (McKhann et al., 1984; Mirra et al., 1991).

CERAD clinical criteria involve the assessment of demographic characteristics and clinical and neurological information. CERAD criteria assess aspects of language, memory and cognitive function. Based on the results of the information provided, CERAD clinical criteria include the diagnosis of possible or probable AD. Utilizing the NINCDS-ADRDA clinical criteria for assessment can result in categories of possible or probable AD (McKhann et al., 1984).

Once a clinical diagnosis of AD is determined, treatments are limited. Since there is currently no known cure for AD, care regimes only involve the management of symptoms. Providing social support and managing symptoms are the primary goals of dementia care (WHO, 2017). Care regimes for dementia emphasize early diagnosis and detecting and treating the associated behavioural and psychological symptoms, as well as supporting care partners (WHO, 2017). To manage individual symptoms associated with irreversible dementias such as AD, both pharmacological (e.g., cholinesterase inhibitors) and non-pharmacological treatments (e.g., counselling) can be utilized (Alzheimer's Association, 2016; Waldemar et al., 2007). Pharmaceutical treatments can be helpful in dealing with the related symptoms. However, only managing the symptoms associated with AD fails to address additional impacts on the individuals themselves as well as their care partners. Combinations of both pharmacological and non-pharmacological approaches have been useful in addressing AD symptoms. However, further research and approaches to treating AD are still needed.

Neuropathologic Criteria

Two of the well-recognized neuropathologic criteria for identifying Alzheimer neuropathology in post-mortem brains include the National Institute on Aging and Reagan Institute (NIA-RI) neuropathologic assessment, and CERAD neuropathologic criteria (Fillenbaum et al., 2008; Hyman et al., 2012).

CERAD neuropathologic criteria were originally developed in 1986 as an approach to provide a standardized protocol for NP scoring based on analysis of neuroimaging techniques (Fillenbaum et al., 2008; Mirra et al., 1991). NPs are counted and labelled based on their frequency in areas of the neocortex. There are four stages that assess plaque density including C0: no NPs, C1: CERAD score sparse, C2: CERAD score moderate, and C3: CERAD score frequent (Fillenbaum et al., 2008; Hyman et al., 2012). Based on an individual's NP score and their age, CERAD criteria suggest a diagnosis of definite, probable, or possible AD (Heyman et al., 2012).

NIA-RI criteria assert that NFTs and NPs are both required for the neuropathologic

diagnosis of AD (Hyman et al., 2012). To assess NFTs, a variety of histochemical stains are required. As criteria for NFTs, the 1997 staging scheme by Braak is utilized in the NIA-RI model (Hyman et al., 2012). Braak proposed six stages of NFT density including Braak Stage 0: no NFTs, Braak Stages I/II: NFTs clustered around the entorhinal cortex, Braak Stages III/IV: NFTs clustered around the hippocampus and amygdala, and Braak Stage V/VI: NFTs distributed throughout the neocortex (Braak & Braak, 1991). To assess NPs, CERAD scoring is also applied. Following this, both the Braak staging and CERAD scores (i.e., NFTs and NP scores) are combined to create an NIA-RI score reflecting the likelihood of AD. The NIA-RI categories include no, low likelihood, intermediate likelihood, and high likelihood (Hyman et al., 2012).

2.1.4 Risk Factors

There are a variety of modifiable and non-modifiable risk factors for AD. Some of the modifiable risk factors relate to behaviours and include smoking, physical activity, diet, alcohol, cognitive training, social engagement, education, sleep, depression, and cardiovascular health (Baumgart et al., 2015; Prince et al., 2014). Non-modifiable risk factors for AD include age, genetic susceptibility (e.g., *APOE*), and family history of dementia (Corrada et al., 2010; Kivipelto & Solomon, 2008; Selkoe, 2001).

Modifiable Risk Factors

Many lifestyle factors influence one's potential risk of developing AD in later life. Smoking is recognized as one of these lifestyle factors. In a study assessing heavy smoking in mid-life, heavy smoking was found to double an individual's risk of developing dementias such as AD in later life (Rusanen et al., 2011). Tobacco use also had a dose-response effect on one's risk of developing AD. Risk of AD in late life was greater in individuals with moderate to heavy smoking levels in comparison to non-smoking individuals (Tyas et al., 2003). Heavy usage and intensity of smoking was also associated with an increase in Alzheimer neuropathology (NPs) (Tyas et al., 2003). Alternatively, quitting smoking was associated with an overall reduced risk of developing dementia (Prince et al., 2014).

Globally, 13.9% of all cases of AD are attributable to smoking, suggesting the strong significance of this modifiable risk factor (Norton et al., 2014).

Another modifiable risk factor for AD is diet. Longitudinal studies that analyzed Mediterranean diets have suggested an association between this type of diet and reduced risk of AD (Morris et al., 2015). Mediterranean diets often include higher intakes of fruits and vegetables, fish and olive oil, while limiting red meat. However, dietary pattern is often also related to other characteristics (e.g., demographic location, exercise, smoking, etc.) (Morris et al., 2015). These other lifestyle characteristics can interfere with the interpretation of risk reduction of AD as a result of diet (Baumgart et al., 2015).

Another relevant modifiable area with respect to AD is vascular health (such as high blood pressure, diabetes and cardiovascular disease). In particular, high blood pressure has been associated with increased risk of heart attacks, heart failure, stroke and kidney disease (Fillit et al., 2008). Hypertension has also been associated with increases in Alzheimer neuropathology and cortical atrophy (the loss of neuronal tissue) as well as risk of developing AD in later life (Fox et al., 1996; Meng et al., 2014). Diabetes, hypertension and related features of the diseases (e.g., increased oxidative stress, high insulin levels) are associated with cerebrovascular disease (Luchsinger et al., 2004). Studies suggest that cerebrovascular disease and diabetes are significantly related to the development of AD and cognitive impairment in later life (Luschinger et al., 2004). Ischemic strokes (caused by impairments in cerebral blood flow) have also been associated with cognitive impairment and dementias in later life. Cerebral infarcts (areas of brain tissue death), which can result from strokes, can cause severe losses in cognitive function and have been associated with AD pathology as well as other subtypes of dementia such as vascular dementia (Grinberg & Heinsen, 2010). Individuals with a history of stroke are approximately 3.5 to 6 times more likely to develop dementia than those who have not had a stroke (Fillit et al., 2008). Other cardiovascular risk factors such as obesity and a lack of physical activity have also been associated with AD.

Physical activity is also a modifiable risk factor since maintaining good physical health has been associated with improvements in cognitive impairment related to AD (Baumgart et al., 2015). In studies assessing the effect of physical activity on cognitive

function in older adults, improvements in cognitive state were seen after only six months (Lautenschlager et al., 2008). Many hypotheses have been put forward as to why physical activity improves cognitive status. However, the most common explanations have been enhancements to brain plasticity (i.e., the ability of the brain to change structure or function in response to stimuli), increased blood flow to relevant brain regions, and increased neurogenesis as a result of physical activity (Kolb & Wishaw, 1998; Lautenschlager et al., 2008). Approximately 13% of all cases of AD are attributable to physical inactivity, further suggesting the importance of remaining physically active (Norton et al., 2014).

Alcohol consumption can also be classified as a modifiable factor influencing risk of AD (Anstey et al., 2009). Some studies suggest that alcohol consumption may decrease susceptibility to cognitive decline associated with AD. In particular, low to moderate levels of alcohol consumption (i.e., ranging from 1 to 14 drinks per week) have been linked to significantly reduced risks of developing AD (Anstey et al., 2009). Other research has found significant increases in risk of AD in individuals with moderate alcohol consumption (drinking 3 to 4 standard glasses per day) (Orgogozo et al., 1997). Heavy alcohol consumption has also been associated with detrimental cognitive changes, such as brain shrinkage (atrophy), neurodegeneration, and cognitive decline (Tyas, 2001). Findings on the association between alcohol consumption and AD may vary as a result of differences in dosage (e.g., amount of alcohol) and limitations in methodologic approaches. The effect of alcohol and the pattern of results is reliant on the amount of alcohol consumed. Differing study designs, the validity of proxy respondents and other methodological challenges may all affect the accuracy of results (Tyas et al., 2000). However, there are many well documented associations between alcohol consumption and risk of morbidity and mortality that suggest the need to critically assess findings with respect to risk of AD (Orgogozo et al., 1997).

Level of educational attainment has also been linked to AD (Baumgart et al., 2015). Level of education is thought to be an indication of intellect as well as level of cognitive stimulation. In a systematic review discussing the relationship between education and AD at the population level, higher educational attainment was consistently associated with significant reductions in the prevalence and incidence of AD (Jefferson et al., 2011; Meng &

D'Arcy, 2012). Conversely, low levels of educational attainment have been associated with an increased risk of developing AD (Meng & D'Arcy, 2012). The most notable attributable factor for AD is low levels of educational attainment, which account for approximately 19.1% of all cases of AD worldwide (Norton et al., 2014). Other factors associated with educational attainment, such as complexity of adult work and regular engagement with intellectually stimulating activities, were also identified as potential protective factors for the onset of AD (Meng & D'Arcy, 2012).

Additional risk factors such as history of depression, moderate or severe traumatic brain injury, and sleep disturbances have all been associated with an increased risk of cognitive decline and dementias such as AD in late life (Baumgart et al., 2015).

Non-Modifiable Risk Factors

Age is the strongest known non-modifiable risk factor for AD (National Institute on Aging, 2016). After the age of 65, the potential risk of developing AD doubles every 5 years, suggesting the notable effect of age as a risk factor (Corrada et al., 2010). Higher likelihood of developing AD in older age is associated with the accumulation of NPs and NFTs as individuals age (Selkoe, 2001). As the population ages, the relevance of this disease will continue to increase.

Genetic and familial factors have also been recognized as non-modifiable risk factors for AD. Familial AD can be inherited through autosomal dominance (Selkoe, 2001). In familial cases, similarities in phenotypes suggest consistent mutations in genes causing AD. Mutations in gene coding can produce greater levels of amyloid beta aggregates causing familial or early-onset AD to occur (Wu et al., 2012). The β-amyloid precursor protein gene is a sequence that assists with making proteins and is found within many tissues and organs in the body (Wu et al., 2012). Mutations in β-amyloid precursor protein (i.e., changes to sequences of exons 16 and 17) and mutations in Presenilin 1 and 2 genes have been associated with excessive production of amyloid β and are a cause of familial AD (Tanzi & Bertram, 2005; Wu et al., 2012).

In sporadic cases of AD, genetics are also an important factor. Apolipoprotein E

(*APOE*) has a major role in lipid metabolism (Kivipelto & Solomon, 2008). *APOE* consists of amino acids and combines with lipids in the body to maintain normal levels of cholesterol (Holtzman et al., 2012). There are three *APOE* alleles: ε2, ε3, and ε4. Inheriting the *APOE*-ε2 allele is thought to be protective against developing AD (Meyer et al., 1998). The inheritance of one or two ε4 alleles has been associated with significant increases in susceptibility and a general predisposition to developing AD (Kivipelto & Solomon, 2008). *APOE*-ε4 has been associated with some of the major features associated with AD pathogenesis such as β-amyloid generation, NFT development and oxidative stress (Kivipelto & Solomon, 2008). The accumulation of β-amyloid peptide has been associated with neuronal death and the progression of symptoms of AD. It is also known that having the genetic risk factor *APOE*-ε4 makes individuals much more likely to develop AD (Kivipelto & Solomon, 2008).

2.2 Cognitive Resilience

2.2.1 Brain and Cognitive Reserve

There are individual differences in level of susceptibility to the brain-related changes and pathology associated with AD. In later life, many older adults have Alzheimer neuropathology present. Clinical expression of dementia and functional impairments can occur when an individual is impacted by brain damage. However, some individuals with Alzheimer neuropathology demonstrate greater reserve and the clinical expression of dementia is less apparent (Stern, 2002). The concept of reserve describes the phenomenon of having Alzheimer neuropathology present and not meeting the clinical diagnosis for AD (Stern, 2002; Stern 2012). Reserve is classified into two different forms: brain reserve and cognitive reserve.

Brain reserve focuses on the quantitative aspects of reserve, such as the number of neurons an individual has or their brain size (Stern, 2012). Brain reserve theories suggest that over time repeated instances of brain damage accumulate (Stern, 2002). However, individuals differ in terms of their brain reserve capacity and ability to cope with the structural changes associated with AD (Stern, 2002). Fewer cases of dementia have been found in individuals with larger brains, suggesting that larger brains may be able to tolerate

more AD neuropathology. Theories on brain reserve suggest that there is a threshold of neuropathologic damage (i.e., functional impairment cut-off) a brain can tolerate before clinical symptoms of AD appear (Stern, 2002; Stern, 2012).

In contrast, cognitive reserve is an active model that describes the phenomenon with respect to brain function rather than brain size (Stern, 2012). In the presence of AD neuropathology, cognitive reserve suggests that the brain actively copes by using pre-existing compensatory approaches (Stern, 2012). Some of the potentially compensatory mechanisms include life experience, levels of intelligence, education and cognitively stimulating activities. Cognitive reserve theory as described by Stern (2002) entails two major forms: reserve and compensation. Neural reserve theory suggests that individual brains are not anatomically different; however, those with greater cognitive reserve have better and more efficient processing mechanisms (Stern, 2002). Individuals with greater reserve have highly efficient neural networks and are better able to maintain function in the presence of AD neuropathology (Stern, 2002). In contrast, neural compensation increases an individual's reserve capacity by recruiting new neural networks for future use (Stern, 2012). In cases where neuropathologic damage occurs, individuals with greater neural compensation would have greater cognitive flexibility to overcome neuropathologic damage (Stern, 2002).

2.2.2 Resilience

Resilience is a dynamic concept that suggests the ability to overcome or adapt in the presence of adversity (Allen et al., 2011). Similarly, resilience is the maintenance of functionality resulting from strength gained throughout life experiences (Allen et al., 2011). The term *cognitive resilience* in this context is used in this thesis to describe the ability of some individuals to maintain cognitive function through both brain reserve and cognitive reserve in the presence of AD neuropathology (Stern, 2012).

One of the most documented examples of cognitive resilience is the association between level of education and AD (Hall et al., 2007). According to the cognitive resilience hypothesis, higher levels of educational achievement should delay the onset of cognitive impairments such as AD. However, once an individual with higher education is diagnosed

with AD, their decline should be more rapid, considering that more pathology would have accumulated prior to the diagnosis (Hall et al., 2007). Examples of cognitive resilience hypotheses are evident in research that has linked low educational attainment to earlier onset of symptoms and diagnoses of AD (Qui et al., 2001). Similarly, combinations of low education and low occupation have been associated with heightened risk for dementia (Stern et al., 1994). Cognitive resilience hypotheses have also been supported by findings of increased rates of decline associated with each additional year of education following diagnosis of AD (Hall et al., 2007). Although the relationship between educational attainment and cognitive resilience has been well documented, the association of cognitive resilience with early-life written language skills, which support educational attainment and similarly reflect intellectual ability, has been less studied.

2.3. Language Skills

2.3.1 Childhood Development of Language Skills

Early-life experiences provide the foundation for later life. Factors from early life have been discussed in literature as influencing language acquisition in childhood. Research has identified a variety of factors such as culture, ethnicity, caregiver age, birth order, television viewing, peers, and quantity and redundancy of child-directed speech as influencing the development of language in children (Hoff, 2006). Language acquisition is an innate and inevitable aspect of humanity. However, the level of environmental support influences individual differences in use of language (e.g., a child's level of complexity in structures, the size of the child's vocabulary) (Hoff, 2006). Aspects of a child's early environment such as the family's socioeconomic status, ethnicity, household literacy rates, parental education or occupation, and household size have all been cited as affecting language acquisition in children (Brewster et al., 2014). Parents can invest in a child's early language environment by providing learning equipment (e.g., books, toys) and investing time on developmental activities (e.g., reading) that improve cognitive and language development (Arif & Albulene, 2016). Findings with respect to children from higher social strata reflect beneficial parental investment since these children were generally able to produce longer

sentences and respond appropriately to adult speech. In standardized tests, toddlers and children (at age five and six years) from higher social strata scored higher and had greater grammatical complexity scores than toddlers and children from lower social strata (Huttenlocher et al., 2002). Parents of low socioeconomic standing were often less able to provide beneficial language environments for their children. Analyses on children from low-income households found that 70% of the children performed below the 50th percentile on sentence complexity scores (Arif & Albulene, 2016; Arriaga et al., 1998). Beneficial household and community socioeconomic measures (e.g., self-perceived childhood and county socioeconomic status, and parental education) have also been related to better cognition in late life (Wilson et al., 2005). However, poor childhood socioeconomic status has not been consistently associated with an increased risk of AD (Wilson et al., 2005).

Other aspects of childhood language development, such as writing, have also been studied. The majority of research in this area focuses on handwriting skills, copying ability, language production, and finger movement (Abbot et al., 2010). However, analyses on children across grade levels provide information on the way in which writing skills change while cognitive and linguistic functions develop (Crossley et al., 2011). When learning to write, a child's neural development plays a key role since it directly corresponds to their ability to process words and utilize aspects of working memory. At second grade, children begin using stronger linguistic connections (e.g., referential pronouns or connectives) and they begin generating better and more cohesive writing (Crossley et al., 2011). By the age of 10, conventional aspects of writing and local coherence begin to emerge and continue to develop until the eighth grade. At the end of high school and into early college years, individual's rate of improvement in written cohesion (i.e., the form, flow and grammatical aspects of sentences) slows down (Berninger et al., 2010). Comparisons between high school students and 3rd year college students have demonstrated no significant differences in terms of unity, organization, development or coherence (Crossley et al., 2011). These analyses further emphasize that developmental patterns were most apparent in children of young ages (i.e., elementary to middle school children). Although many studies discuss childhood writing development, few results have been linked to later-life cognitive outcomes. However,

research on other aspects of language throughout the life course can provide additional clues for preventing or delaying the onset of dementia.

2.3.2 Multilingualism

Another major area of language research that has been commonly studied with respect to late-life cognitive impairment is multilingualism. This area of language research focuses on the relationship of number of languages spoken and understood with cognitive outcomes such as dementia. However, findings have been inconsistent. In clinic-based studies, multilingualism has been shown to delay the onset of cognitive impairment, providing a protective effect. For example, when comparing monolingual individuals to bilingual individuals at a memory clinic, Bialystok et al. (2007) found that dementia was delayed by four years in bilingual individuals. Similar research by Craik et al. (2010) found that among individuals with probable AD, bilinguals reported symptoms significantly later than monolinguals (mean=5.1 years). In contrast, population-based studies typically have shown no significant association between multilingualism and dementia (e.g., Brewster et al. 2014; Perquin et al., 2013; Yeung et al., 2014). Methodological issues, such as selection biases, information biases and confounding, may contribute to these inconsistent findings. Alternate measures of language skills may prove more consistently related to cognitive status in later life and add to our understanding of the association between early-life language skills and late-life cognitive impairment.

2.3.3 Written Language Skills

2.3.3.1 Oral Versus Written Language Skills

While multilingualism is the predominant metric of language skills in current research on language and dementia, alternative indicators include oral and written language measures. Early comparisons between oral and written language samples suggested that written language was more efficient than oral language when comparing amount of information conveyed in comparison to the number of words in a sentence (e.g., in the number and use of such constituents as prepositional phrases and nominalizations (adjectives

or verbs morphologically converted into nouns) (Chafe & Danielewicz, 1987). However, to increase understanding of the appropriateness of written and oral language samples, recent studies have reassessed these comparisons.

Mitzner and Kemper (2003) compared written and oral autobiographical narratives collected from 118 participants in the Nun Study, a longitudinal study of aging with access to archives containing narratives written in early adulthood. Similar research by Mueller et al. (2015) focused on verbal fluency in oral language in a longitudinal study of individuals with a family history of AD. Mitzner and Kemper (2003) used propositional density (i.e., idea density) measured per 10 words and grammatical complexity in their analysis of narratives. In contrast, Mueller et al. (2015) used measures of verbal fluency (e.g., phonemic fluency) and category fluency (e.g., semantic fluency that measures knowledge and memory) from samples of individuals with mild cognitive impairment (MCI) and individuals who were cognitively healthy. Fluency tests require participants to list as many words as they possibly can within a set time period. Phonemic fluency requires an individual to list words that begin with a specific letter, whereas semantic tests relate to a category (e.g., vegetables, fruits) (Mueller et al., 2015).

When comparing written and oral narratives, written samples had many advantages. Oral narratives tended to be longer than written language samples (Ravid & Berman, 2006). However, in written language samples, greater control over word usage and the ability to express more information was evident (i.e., significantly higher levels of idea density) in comparison to oral language samples ($p<0.01$) (Mitzner & Kemper, 2003). Written language samples contained significantly more diverse vocabulary than oral samples ($p<0.05$) and lower proportions of sentence fragments ($p<0.01$). Oral and written language samples, however, did not vary significantly in developmental level (grammatical complexity), suggesting that written language samples were no more grammatically complex than oral language samples (Mitzner & Kemper, 2003).

Studies comparing written and oral language tended to be longitudinal in nature. Mitzner and Kemper (2003) utilized a homogeneous sample (the Nun Study), which included individuals who were well educated, lived together and shared in many daily activities. The

sample used by Mueller et al. (2015) was a cohort of individuals from a Wisconsin town with a higher than average baseline IQ, and participants were selected based on their age and family history of AD. Individuals with prior neurological conditions at baseline were also excluded from analyses. However, individuals with other emerging conditions that could influence cognition were not excluded from the study, and thus comorbid conditions (e.g., hypertension, metabolic syndrome) could have affected study findings (Mueller et al., 2015). These samples had limited generalizability, with the possibility of cohort effects due to shared temporal experiences (e.g., birth year, common life experiences). Other limitations include the inability of many participants who had dementia to produce both an oral and written language sample (Mitzner & Kemper, 2003). This may have influenced conclusions since those with more severe dementia would have been less likely to have provided both oral and written language samples and been included in the study.

Idea density from written language samples was associated with cognitive status as measured by the Mini-Mental State Examination (MMSE), a neuropsychological test of cognitive function including multiple cognitive domains. In addition, individuals with limited vocabularies performed worse on the MMSE (Mitzner & Kemper, 2003). Similarly, decreases in other measures from written language samples (e.g., verbal fluency and switching) were associated with cognitive impairments such as MCI (Mueller et al., 2015). Written language samples were stronger than oral samples at predicting cognitive status and later-life cognitive ability in older adults (Mitzner & Kemper, 2003), supporting preference for written rather than oral language samples in studies of cognition.

Written language skills can be understood as an indication of an individual's ability to process, encode, organize, and retrieve information. These essential skills relate to educational attainment and success throughout life (Jefferson et al., 2011; Snowdon et al., 1996). Linguistic ability has been represented in research as complexity in terms of information conveyed (e.g., measures such as idea density and mean length of utterance), and in grammatical construction (grammatical complexity or Developmental Level) (Cheung & Kemper, 1992; Snowdon et al., 1996). Although linguistic complexity metrics vary, they

provide important measures for researchers to better examine developmental trends (e.g., childhood grammatical skills) and fluctuations in cognitive states (Cheung & Kemper, 1992). Two major indicators of written language skills, idea density and grammatical complexity, are commonly discussed in existing literature.

2.3.3.2 Idea Density and Grammatical Complexity

Idea density (originally known as propositional density) describes the number of ideas expressed per number of words and has been linked to quality of writing, processing efficiency, and vocabulary (Brown et al., 2008; Cheung & Kemper, 1992). Idea density is also a measure of information load or semantic content of written language (Cheung & Kemper, 1992). The concept of ideas or idea units relates to measuring the usage of elementary propositions that include verbs, adjectives or prepositional phrases (Riley et al., 2005). Propositions (an underlying measure of ideas) include three different categories: statements that indicate expressive actions or states, statements that express restrictions or limitations, and statements that suggest connections (e.g., expressing causality or contrast events) (Cheung & Kemper, 1992). Idea density was originally conceptualized by Kintsch and Keenan (1973), who compared the amount of information to the number of words per sentence. Idea density can also be used as an indicator of an individual's ability to define and process words. An example of how idea density is scored is provided in Appendix C.

Idea density has been shown to decline parallel to declines in performance in semantic, episodic and spatial ability indexes, suggesting a connection between idea density and memory (Farias et al., 2012). Idea density has been strongly associated with vocabulary size (Kemper et al., 2001b). Idea density has not, however, been associated with executive function (Farias et al., 2012).

Grammatical complexity (originally based on the Developmental Level Metric) is a measure of the complexity of grammar in sentences. Grammatical complexity involves the rating of sentences from 0 (simple sentences) to 7 (complex sentences) (Snowdon et al., 2000). A full description of each level of grammatical complexity is listed in Appendix C. Grammatical complexity is a measure of embedding (i.e., when a clause or concept is

included or embedded within another clause), and has been associated with working memory, which is an aspect of executive function (Cheung & Kemper, 1992). Executive function is a general system of processes that include the management, regulation, and coordination of attention and control. Executive function also involves the modification of behaviour as responses to environmental and other changes (Carpenter et al., 2000). Using measures that assess working memory (e.g., digit span and reading span), associations between grammatical complexity and executive function have been found. Digit span describes the ability to retain digits (whether forward or backward). Similarly, reading span describes the ability to retain sentences that were read (Kemper et al., 2001). Findings have highlighted the relationship between grammatical complexity and span measures, with individuals with higher grammatical complexity more likely to have higher digit span and reading span scores (Kemper et al., 2001).

2.3.3.3 Written Language Skills and Cognition

Measures of early-life written language abilities have been analyzed with respect to later-life cognitive function. According to research by Snowdon et al. (1996) in the Nun Study, low idea density significantly increased the risk of poor cognitive function and AD in late life (Snowdon et al., 1996). Idea density was also strongly related to later-life neuropathology. Those with low idea density tended to have lower brain weights (i.e., less than 1000 grams) with higher odds of moderate or severe cerebral atrophy (OR=4.7; 95% CI: 1.1-20.0) (Riley et al., 2005); they were also more likely to meet neuropathologic criteria for AD (OR=4.9; 95% CI: 4.6-5.3) (Snowdon et al., 2000). Linguistic measures, such as proportion of fragments and idea density from written language samples, have also been significantly associated with MMSE scores (Mitzner & Kemper, 2003). Other measures of language skills such as generative semantic verbal fluency tests (measuring phonemic fluency and semantic fluency) were also used to measure declines in older adults developing AD (Pakhomov & Hemmy, 2014). Semantic performance (i.e., generating words that belong to a specific semantic category) was significantly associated with risk of developing dementia ($p<0.01$) but not memory impairment (Pakhomov & Hemmy, 2014). These

findings suggest the potential value of assessing the impact of linguistic ability in early life on additional cognitive outcomes, such as cognitive resilience.

There were also other important variables to consider when analyzing the association between written language skills and cognitive status. Many written linguistic measures were significantly associated with education. Individuals with higher levels of education used more clauses, produced longer sentences and had fewer fragments in their written language when compared to their less educated counterparts (Mitzner & Kemper, 2003). Age was not significantly associated with idea density (Mitzner & Kemper, 2003), but was negatively associated with working memory, suggesting a possible parallel decline in associated measures (e.g., grammatical complexity) (Cheung & Kemper, 1992). The association of language skills (e.g., verbal fluency) and risk of MCI was not significantly influenced by genetic factors such as *APOE* or family history of AD (Mueller et al., 2014).

Although previous research has suggested an association between weak early-life written language skills and poor cognitive function in later life, some limitations reduce the strength of these findings. Small sample sizes were a potential limitation; for example, Cheung and Kemper (1992) analyzed only 30 language samples in research assessing early-life written language skills. Criteria for assessing AD neuropathology relied only on Braak staging or CERAD criteria, meaning that combinations of both NFTs and NPs (i.e., through use of NIA-RI criteria for AD) were not investigated (Pakhomov & Hemmy, 2014; Riley et al., 2005). Neuropsychological tests at baseline did not include tests of executive function, meaning cognitive impairments in this cognitive domain could have been present at baseline (Pakhomov & Hemmy, 2014). Prior studies only demonstrated associations between idea density and cognitive outcomes; grammatical complexity has not been as frequently investigated.

2.3.4 Summary of Written Language Skills

Literature on written language skills suggests a variety of potential opportunities for further exploration. When comparing the strength of written and oral language samples, studies have suggested that written language contained significantly higher levels of idea

density (Mitzner & Kemper, 2003). Overall, written language tended to involve more diverse vocabulary and information and involved greater use of expression when compared to oral language (Mitzner & Kemper, 2003). However, no significant differences were determined between written and oral language in terms of grammatical complexity. Written language however, did show stronger power with respect to differentiating between cognitive status and ability levels in older adults (Mitzner & Kemper, 2003). These findings suggested the appropriateness of using written samples over oral samples as an indication of language skills.

Measures of written language skills (predominantly idea density) were associated with poor cognitive function and AD in later life. Strong associations of low idea density with poor cognitive function or the progression of AD in later life were identified in the Nun Study (Riley et al., 2005). However, small sample sizes were common, potentially limiting the power and ability to detect the true effect of written language skills. Additional limitations included methodologic concerns with measurement or sample. For example, much of the existing literature only relied on one form of AD neuropathologic criteria (e.g., Braak staging or CERAD), meaning that assessments of both NFTs and NPs were not available. In other studies, only participants with a family history of AD were selected (Mueller et al., 2015). These limitations reduced the ability of these studies to fully evaluate the potential association between language skills and cognitive impairment.

Chapter 3
Study Rationale and Research Questions

3.1 Study Rationale

Although early-life written language has been analyzed with respect to dementia, gaps still exist. Previous research using Nun Study data assessed the association between language skills (i.e., idea density and grammatical complexity) and dementia. However, this research was conducted on an earlier, smaller sample, which limited the ability to broaden study findings and assess other cognitive outcomes. Previous Nun Study research has not assessed the association between written language skills (idea density and grammatical complexity) and later-life cognitive resilience incorporating aspects of both cognitive and brain reserve. In work by Snowdon et al. (1996), it was predicted that higher linguistic ability could be a potential indicator of cognitive resilience. However, only one individual from their sample population met the neuropathologic criteria for AD and did not show the expected cognitive decline (i.e., remained cognitively resilient) (Snowdon et al., 1996). While cognitive resilience could not be examined at that time, additional years of follow-up in the Nun Study has expanded the sample available for assessment of cognitive resilience. The hypothetical construct of cognitive resilience provides new opportunities to further explore the impact of early-life written language skills on later-life cognitive resilience while addressing some of the limitations of previous studies.

Furthermore, existing literature on written language skills has not assessed cognition using neuropathologic criteria that assess both NFTs and NPs. The present research aimed to further research in this area by analyzing the association between written language skills and cognitive resilience using criteria that assess both NFTs and NPs.

The aims of this research were: (1) to characterize the distribution of cognitive resilience with respect to early-life written language skills (idea density and grammatical complexity), (2) to determine the relationship between early-life written language skills and cognitive resilience, (3) to determine whether this association holds when adjusted for indicators of cognitive reserve, and (4) to determine whether the association between early-

life written language skills and cognitive resilience holds when adjusted for indicators of brain reserve.

The present research utilized longitudinal data from the Nun Study. Univariate, bivariate and multivariable analyses were conducted to assess the association between written language skills and cognitive resilience. Measures of written language skills (idea density and grammatical complexity) were obtained from autobiographies previously coded for these measures. The outcome, cognitive resilience, was measured based on data from late-life cognitive assessments and post-mortem neuropathologic assessments (NIA-RI and CERAD criteria). Analyses included relevant indicators of cognitive reserve (education) and brain reserve (cortical atrophy and cerebral infarcts).

Further understanding of the influence of early-life written language skills on building cognitive resilience could be potentially helpful in prioritizing interventions in early life, such as focusing on strategies that improve working memory in schools. Additionally, developments in our understanding of interactions between written language skills and other covariates throughout the life course may provide beneficial clues into how other characteristics influence which individuals develop resilience.

3.2 Research Questions

1. How does cognitive resilience vary by level of early-life written language skills (idea density and grammatical complexity)?
2. Are early-life written language skills associated with cognitive resilience, and does this association hold after adjusting for standard confounders (age and *APOE)*?
3. Does the association between early-life written language skills and later-life cognitive resilience hold after adjusting for indicators of cognitive reserve (i.e., education)?
4. Does the association between early-life written language skills and later-life cognitive resilience hold after adjusting for indicators of brain reserve (i.e., cortical atrophy and brain infarcts)?

Based on previous literature, it was hypothesized that: (1) individuals with weak early-life written language skills would be less likely to be cognitively resilient in later life, (2) the association between weak written language skills and cognitive resilience would vary by levels of cognitive reserve (using education as an indicator), and (3) the association between weak written language skills and cognitive resilience would vary by levels of brain reserve (using cortical atrophy and cerebral infarcts as indicators).

Chapter 4
Methodology

4.1 Literature Search Strategy

The purpose of this systematic search was to identify literature related to early-life written language skills and cognitive resilience. A systematic search of relevant literature was conducted in June 2017 using PubMed and PsycINFO. The PubMed database included articles that were published from 1950 onwards. The full literature search template can be found in Appendix A. The literature search was restricted to peer-reviewed articles in English. Search terms for early life and child, written language skills, age and cognitive resilience were considered. For this search strategy, concepts related to "early life [tiab]" or "adolescent [MeSH]" or "child*[tiab]" and "handwriting [tiab]" or "language [MeSH]" and "older adult [tiab]" and "cognitive reserve [MeSH]" or "Alzheimer disease [MeSH]" or "dementia [MeSH]" or "cognitive decline [all fields]" were included. The PubMed search retrieved 341 articles before further screening.

A second systematic search was first conducted using PsycINFO in November 2016 and repeated in June 2017. The PsycINFO database included publications from after 1840. Articles for this search were restricted to peer-reviewed articles in English. The full literature search template can be found in Appendix A. PsycINFO index terms related to "early experience" and "childhood development" and "written communication" and "aging" or "Alzheimer's disease" were used. The PsycINFO literature search retrieved 201 records prior to screening.

Articles retrieved from both databases, PubMed and PsycINFO, were excluded if the exposure was not related to written or oral language measures or if relevant cognitive states (e.g., cognitive resilience, dementia, AD or cognitive decline) were not discussed. After exclusion criteria were applied, a total of 16 articles from PubMed, and 9 articles from PsycINFO were retrieved. Following the removal of duplicate articles, a total of 10 articles remained. Summaries of these reviewed articles can be found in Appendix B.

4.2 Data Source: The Nun Study

4.2.1 Background

The Nun Study is a longitudinal study focused on aging among members of a religious congregation, the School Sisters of Notre Dame, in the United States (Snowdon et al., 1997). In Mankato, Minnesota, Dr. David Snowdon began a pilot study on a small group of religious sisters from the School Sisters of Notre Dame in 1986. Initial pilot studies led to the present Nun Study, which expanded to include members from religious provinces including the Midwestern, southern and eastern regions of the United States (Snowdon et al., 1997).

Nun Study participants had similar lifestyles and living circumstances such as level of social support, access to medical services, marital status, and tobacco and alcohol usage. However, some individual characteristics may have varied such as personal dietary choices and comorbid diseases. The majority of the religious sisters worked as teachers, while the remaining were house sisters and were responsible for household work in the convent. The relative homogeneity of the Nun Study population is advantageous in reducing the effects of confounding variables.

4.2.2 Participants

Members of the School Sisters of Notre Dame who were aged 75 and older at baseline were asked to participate in the Nun Study (Snowdon et al., 1997). Out of a possible 1,031 eligible women, 678 (66 percent) agreed to participate in the study. When the study began in 1991, participants were between the ages of 75 and 102, with an average age of 83 (Danner et al., 2001). All of those who agreed to participate in the study consented to access to archival records, as well as physical and cognitive assessments. Participating sisters also agreed to donate their brains following death (Snowdon, 1997). Non-participants were similar to participants in terms of age, place of origin, race, and mortality rates (Snowdon et al., 1996).

4.2.3 Autobiographies

Beginning in 1930, autobiographies were required to be written upon formal entry into the religious congregation of the School Sisters of Notre Dame (Danner et al., 2001). Individuals were instructed to write autobiographical sketches of their life including events that led them to the convent (Snowdon et al., 2000). Other components that autobiographies included were the individual's place of birth, interesting events from their childhood, and an indication of their religious lifestyle (Patzwald & Wildt, 2004) Autobiographies were not to exceed two to three hundred words and were to be written on a single sheet of paper, suggesting that the autobiographies were generally consistent in terms of length (Danner et al., 2001). The autobiographies included in the Nun Study were written between the years of 1931-43 prior to the sisters formally joining the convents (Riley et al., 2005). Autobiographies were selected for scoring or further analysis if they were handwritten, and the participant was proficient in English and born in the United States (Danner et al., 2001). The autobiographies that met these conditions were from two convents (Milwaukee, Wisconsin and Baltimore, Maryland). In total, 180 autobiographies were scored for written language skills (idea density and grammatical complexity) (Danner et al., 2001).

4.3 Analytic Sample

To assess the relationship between early-life written language skills and cognitive resilience, the analytic sample was restricted to participants with scored hand-written autobiographies (n=180) (see Figure 1). Furthermore, participants who were still alive (n = 29) when the data set was assembled were excluded since complete neuropathological assessments based on the CERAD and NIA-RI criteria were required for the definition of cognitive resilience. In addition, those who did not meet CERAD (i.e., yes/definite or probable) or NIA-RI neuropathologic criteria (i.e., yes, intermediate likelihood, or high likelihood) were excluded as they did not have the neuropathologic basis for cognitive resilience. Participants were also excluded if they were missing data for key variables. Based on these restrictions and exclusions, the final analytic samples for cognitive resilience based on CERAD (n=56) and NIA-RI (n=42) criteria were determined. Tables

assessing selection bias in the analytic samples for either CERAD and NIA-RI are provided in Appendix E.

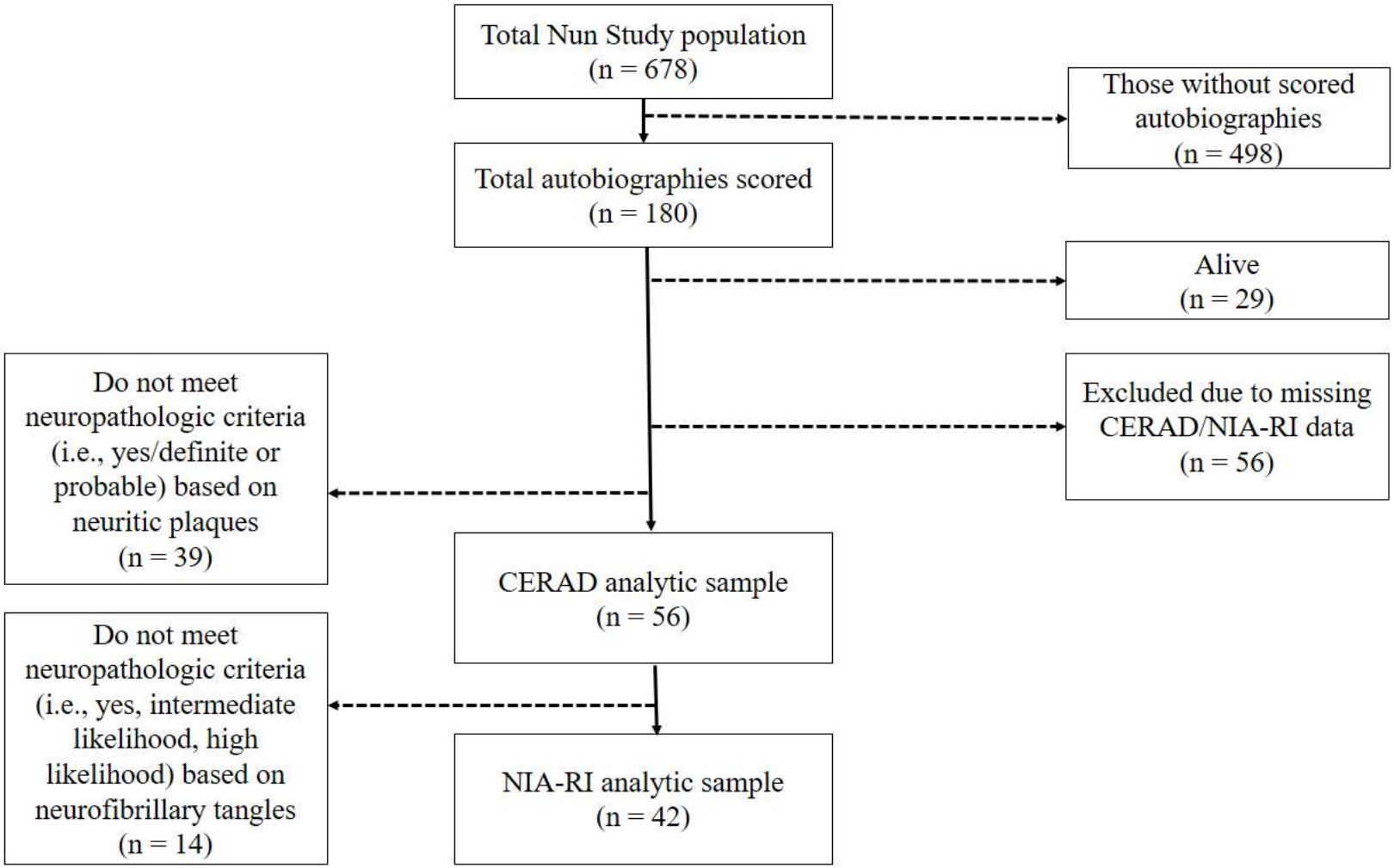


Figure 1. Flowchart of Analytic Sample

Abbreviations: CERAD = Consortium to Establish a Registry for Alzheimer's Disease; **NIA-RI** = National Institute of Aging – Reagan Institute

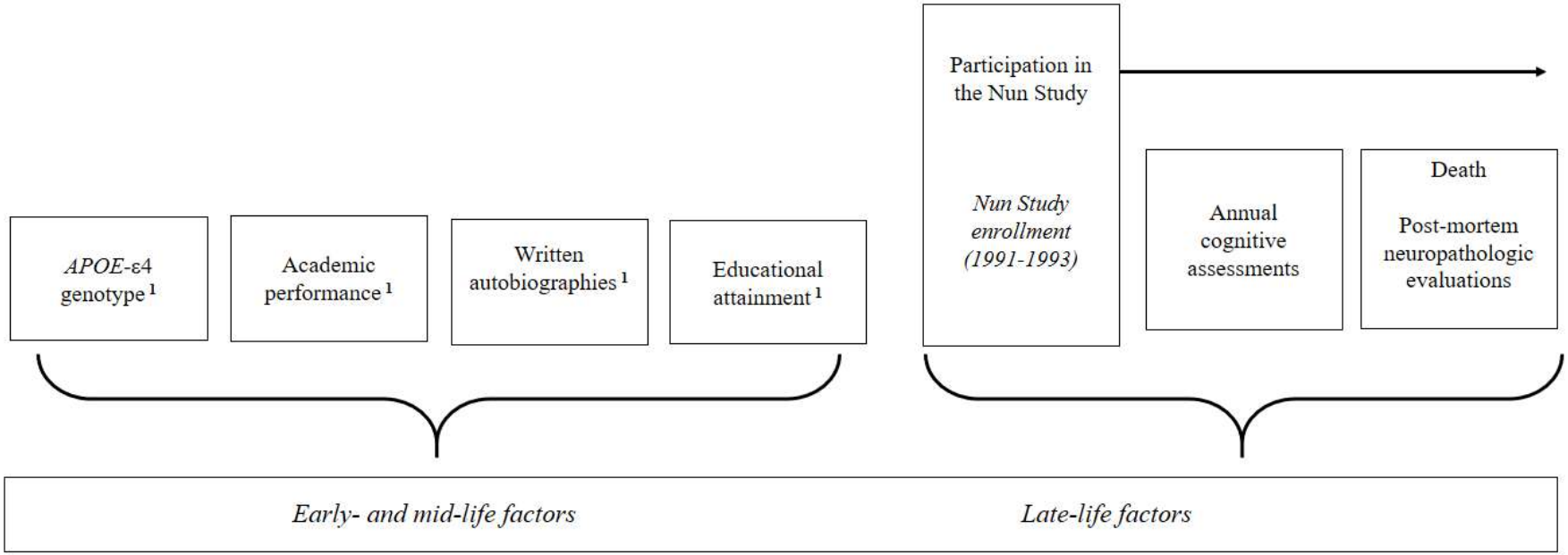
4.4 Measures

Factors from throughout the life course were relevant in the present study. Figure 2 below provides a timeline of these key measures.

4.4.1 Exposure Measures

Measures of written language skills were assessed from the last ten sentences of each handwritten autobiography in the Nun Study by a single linguistic coder (Snowdon et al., 2000). Ten autobiographies were scored independently by an additional coder; the two coders had a high level of agreement and reliability (Snowdon et al., 2000). Idea density was defined as the average number of ideas expressed per ten words (Kemper et al., 2001b). Complex propositions and other relationships between ideas were also counted. Complex propositions include inferred causality or contain temporal components (Riley et al., 2005). Grammatical complexity scores were determined using the Developmental Level metric, which classifies sentences ranging from 0 (simple sentences) to 7 (complex sentences) in terms of grammatical complexity (Snowdon et al., 2000). A description of the calculation of idea density and grammatical complexity is provided in Appendix C.

Scores for both idea density and grammatical complexity were calculated separately and ranked for each convent (Milwaukee or Baltimore) because of variations in the distribution of these scores between the two convents (Riley et al., 2005). Idea density and grammatical complexity scores were divided into quartiles as these measures were not normally distributed. For research question one, analyses of the four quartiles for both idea density and grammatical complexity are provided. For the following research questions, the top three quartiles of both idea density and grammatical complexity were collapsed and classified as “higher”, and the bottom quartile was defined as “low”. Both idea density and grammatical complexity were collapsed into “low” and “higher” groups because small sample sizes required combined categories and significant differences in cognitive resilience were noted for the lowest quartile when compared to the top three quartiles. These categories have also been used in previous studies of these exposures in the Nun Study (Snowdon et al., 1996; Snowdon et al., 2000).



[1] Academic performance (as represented by high school grades), written autobiographies, and educational attainment reflect exposures that occurred prior to Nun Study enrollment. Data on early- and mid-life exposures (e.g., *APOE*- ε4), were collected later, and other exposures were obtained from archival records during the Nun Study (e.g., written autobiographies).

Figure 2. Timeline of Relevant Nun Study Measures

4.4.2 Outcome Measures

The outcome measure, cognitive resilience, was defined as avoiding the clinical symptoms of dementia despite having Alzheimer neuropathology present. This definition was adapted from Stern's (2002) description of cognitive reserve. Cognitive resilience was operationalized as not meeting the clinical diagnosis of dementia according to DSM-IV criteria at last assessment while fulfilling CERAD neuropathologic criteria ("definite" or "probable") or NIA-RI neuropathologic criteria ("definite", "intermediate" or "high" likelihood) for AD (Hyman et al., 2012).

The duration or the time frame between the last cognitive assessment and death at the individual level was considered in the operationalization of cognitive resilience. In the CERAD sample (n=56), the shortest duration was a month and a half with the maximum duration being almost two years (1.89 years). The median duration in the CERAD sample was approximately half a year (0.51 years). In comparison, in the NIA-RI sample (n=42), the shortest duration between the last cognitive assessment and death was a month and a half (0.04 years), and the longest duration was slightly over a year and a half (1.65 years). The median of the NIA-RI sample approximately six months (0.53 years). Using the combined sample of meeting either CERAD or NIA-RI criteria (n=62), the mean duration was approximately seven months (0.61 years) with the median being about six months (0.51 years). The smallest duration between last cognitive assessment and death in the combined sample was a month and a half (0.04 years) and the largest was close to a year and a half (1.37 years).

4.4.3 Covariates

In the Nun Study, potential confounding by many variables was addressed by the homogeneity of the sample (see Section 4.2.1). Standard confounders included *APOE* and age. To understand the association between written language skills and late-life cognitive resilience, additional variables reflecting both cognitive reserve and brain reserve needed to be considered (see Section 2.2.1). Level of education was considered an indicator of cognitive reserve and was adjusted for in relevant multivariable regression models. Variables that were indicators of brain reserve (such as presence of cortical atrophy or presence or number of cerebral infarcts) were also addressed through adjustment in multivariable regression models.

4.5 Analysis

Analyses were conducted using SAS 9.4 statistical software (SAS Institute Inc., Cary, North Carolina).

4.5.1 Descriptive Analyses

Descriptive analyses were conducted on all relevant exposures (idea density and grammatical complexity) and outcomes (cognitive resilience). Analyses included descriptive statistics (e.g., means, standard deviations), univariate and bivariate techniques. Descriptive techniques were used to summarize participant characteristics (Tables 1a and 1b) and to characterize cognitive resilience by level of written language skill for research question one (Tables 2a, 2b, 2c and 2d).

When analyzing the descriptive tables, Pearson chi-square tests were used to compare two categorical variables to determine whether significant associations existed between the variables of interest. T-tests were also utilized when comparing the mean of a continuous variable across two subgroups of a dichotomous categorical variable. In cases where the variances were unequal, rather than using pooled variance, the Satterthwaite method was used.

4.5.2 Multivariable analyses

Logistic regression models were used to address the second and third research questions. First-order interactions between exposures and covariates were tested, and when significant interactions were found, models were stratified by level of the covariate. When models failed to run due to quasi-complete separation of data points, Firth logistic regression techniques were used. Firth regression is a penalized likelihood approach used to provide parameter estimates when there are issues of separability (SAS Insitute Inc., 2017b). Firth regression is advantageous for analysis of small samples and requires less computational memory resources to complete in comparison to Exact regression (SAS Institute Inc., 2017b).

The Hosmer-Lemeshow goodness of fit test was used to assess the fit of models. In addition, residual diagnostic testing and assessment of influential outliers were completed using the INFLUENCE command under PROC LOGISTIC (SAS Institute Inc., 2017). Using the INFLUENCE subcommand, DFBETA, C and CBAR output were produced. These statistics provided the standardized difference for each parameter estimate, in addition to measurements of confidence interval displacements (Kleinbaum et al., 2013). Based on a value of ± 1.96 (i.e., a

significance level of $p<.05$), no observations were considered influential outliers (Kleinbaum et al., 2013). Multicollinearity of exposures and covariates using the variance inflation factor was also assessed (Farrar & Glauber, 1967). A model was considered to be affected by multicollinearity if it the variance inflation factor was ≥10 (Farrar & Glauber, 1967). However, in final models no significant evidence of multicollinearity was determined.

4.6 Ethics

Ethics clearance for the Nun Study was originally obtained in 1990 from the University of Kentucky. The Nun Study later shifted its base to the University of Minnesota. This research was consistent with the scope of a grant that has been approved by the Office of Research Ethics at the University of Waterloo (ORE #20174).

To ensure confidentiality, study participants were identified by participant ID numbers rather than by name. Nun Study data are stored at the University of Waterloo on a password-protected research server. Researchers working with Nun Study data also sign a statement indicating their agreement to follow requirements for working with confidential data.

Chapter 5
Results

5.1 Descriptive Analyses

Tables 1a and 1b summarize the descriptive characteristics of the analytic sample and the distribution of written language skills with respect to the outcome, cognitive resilience, defined according to CERAD (n=56) and NIA-RI (n=42) neuropathologic criteria. In the CERAD sample (Table 1a), the mean age at death was 89.3 years, and 48% (n=27) of participants showed cognitive resilience. In the NIA-RI sample (Table 1b), the mean age at death was similar (89.2 years). However, 31% of individuals in this analytic sample (n=13) showed cognitive resilience.

The descriptive analyses highlight that both analytic samples were highly educated. In the CERAD sample, only 9% of participants had a high school level of education or less. The NIA-RI sample had similar levels of education, with only 9.5% of participants having high school level of education or less.

In both samples, the majority of participants did not have *APOE*-ε4 alleles present. A total of 66.1% of participants in the CERAD analytic sample and 57.1% of individuals from the NIA-RI sample did not possess any *APOE*-ε4 alleles.

For the early-life written language skills variables (idea density and grammatical complexity), the overall variable (i.e., all four quartiles), and low written language skills (i.e., two level variable: lowest quartile versus higher three quartiles) were used in the analyses. In the CERAD sample, chi-square tests identified that overall idea density was significantly associated with cognitive resilience (p=0.048). Low idea density was also significantly associated with the outcome (p=0.01). Chi-square analyses revealed that overall grammatical complexity (p=0.04) and low grammatical complexity (p=0.007) were both significantly associated with cognitive resilience.

In the NIA-RI sample, bivariate analyses revealed that only idea density was significantly associated with having cognitive resilience in later life. Overall idea density (p=0.001) and low idea density (p=0.0005) were both strongly associated with cognitive resilience. However, overall grammatical complexity (p=0.30) and low grammatical complexity (p=0.09) were not significantly associated with cognitive resilience in bivariate analyses of the NIA-RI analytic sample.

Table 1a. Participant Characteristics by Cognitive Resilience Status, CERAD Criteria (n=56)

			Cognitive Resilience (CERAD Criteria)	
Characteristic		Total (n=56)	Yes (n=27)	No (n=29)
Age at death, mean (SD)		89.29 (3.06)	89.35 (3.31)	89.23 (2.85)
Level of education (%)	<High school	3.57	0.00	6.90
	High school	5.36	0.00	10.34
	Bachelor's degree	42.86	51.85	34.48
	≥ Master's degree	48.21	48.15	48.28
Presence of *APOE*-ε4 (%)	No	66.07	77.78	55.17
	Yes	33.93	22.22	44.83
Idea density* (%)	Quartile 1	23.21	7.41	37.93
	Quartile 2	21.43	25.93	17.24
	Quartile 3	21.43	22.22	20.69
	Quartile 4	33.93	44.44	24.14
Idea density* (low vs. higher, %)				
	Quartile 1 (low)	23.21	7.41	37.93
	Quartiles 2-4 (higher)	76.79	92.59	62.07
Grammatical complexity* (%)	Quartile 1	28.57	11.11	44.83
	Quartile 2	17.86	25.93	10.34
	Quartile 3	23.21	29.63	17.24
	Quartile 4	30.36	33.33	27.59
Grammatical complexity** (low vs. higher, %)				
	Quartile 1 (low)	28.57	11.11	44.83
	Quartiles 2-4 (higher)	71.43	88.89	55.17

* $p<.05$; ** $p<.01$

Abbreviations: ***APOE*-ε4** = Apolipoprotein-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **SD** = Standard deviation

Table 1b. Participant Characteristics by Cognitive Resilience Status, NIA-RI Criteria (n=42)

			Cognitive Resilience (NIA-RI Criteria)	
Characteristic		Total (n=42)	Yes (n=13)	No (n=29)
Age at death, mean (SD)		89.19 (3.39)	90.38 (2.61)	88.66 (3.60)
Level of education (%)	<High school	4.76	0.00	6.90
	High school	4.76	0.00	6.90
	Bachelor's degree	45.24	61.54	37.93
	≥ Master's degree	45.24	38.46	48.28
Presence of *APOE*-ε4 (%)	No	57.14	69.23	51.72
	Yes	42.86	30.77	48.28
Idea density** (%)	Quartile 1	38.10	0.00	55.17
	Quartile 2	19.05	38.46	10.34
	Quartile 3	14.29	15.38	13.79
	Quartile 4	28.57	46.15	20.69
Idea density** (low vs. higher, %)				
	Quartile 1 (low)	38.10	0.00	55.17
	Quartiles 2-4 (higher)	61.90	100.00	44.83
Grammatical complexity (%)	Quartile 1	35.71	15.38	44.83
	Quartile 2	14.29	23.08	10.34
	Quartile 3	26.19	30.77	24.14
	Quartile 4	23.81	30.77	20.69
Grammatical complexity (low vs. higher, %)				
	Quartile 1 (low)	35.71	15.38	44.83
	Quartiles 2-4 (higher)	64.29	84.62	55.17

* p<.05; ** p<.01

Abbreviations: ***APOE*-ε4** = Apolipoprotein-ε4; **NIA-RI** = National Institute of Aging – Reagan Institute; **SD** = Standard deviation

5.2 Multivariable Analyses

5.2.1 Research Question 1: How does cognitive resilience vary by level of early-life written language skills (idea density and grammatical complexity)?

Tables 2a through to 2d summarize how cognitive resilience varies by level of idea density and grammatical complexity. Tables 2a and 2b present participant characteristics stratified by level (lowest quartile versus higher) of idea density and grammatical complexity with respect to cognitive resilience sample based on CERAD criteria, whereas Tables 2c and 2d summarize results for the NIA-RI sample.

In Table 2a, only 15% of individuals with low idea density showed cognitive resilience. Similar findings were present in the low grammatical complexity group where only 19% of individuals showed cognitive resilience. Of those who had cognitive resilience in the low idea density category, 100% had a Master's degree or higher. In the low grammatical complexity group, 100% had at least a Bachelor's degree. In both low strata of written language skills (i.e., idea density and grammatical complexity), all individuals who had cognitive resilience had no *APOE-ε4* alleles.

Of those with higher written language skills (CERAD sample), individuals who showed cognitive resilience tended to have higher levels of education (i.e., Bachelor's degree or higher). There were no major differences in terms of *APOE-ε4* status between those with or without cognitive resilience among those with higher idea density. However, for higher grammatical complexity, individuals who had cognitive resilience were less likely to have *APOE-ε4* alleles.

Tables 2c and 2d present participant characteristics by level (lowest quartile versus higher 3 quartiles) of written language skills and cognitive resilience status based on the NIA-RI sample. For the lowest quartile of idea density, there were no individuals with cognitive resilience. In the lowest quartile of grammatical complexity, only 13% of individuals showed cognitive resilience and of these individuals, all had a bachelor's degree as their highest level of education and none were *APOE-ε4* carriers. Individuals with higher levels of written language skills also tended to have higher levels of education since all individuals with cognitive resilience had a Bachelor's degree or higher. Overall idea density (quartiles 1-4) and low idea density (lowest quartile versus higher three quartiles) were both significantly associated with cognitive resilience ($p<.05$) in the higher grammatical complexity group.

Table 2a. Participant Characteristics Stratified by Idea Density, CERAD Criteria (n=56)

	Idea Density					
	Low			Higher		
	Cognitive Resilience (CERAD Criteria)			Cognitive Resilience (CERAD Criteria)		
	Total (n=13)	Yes (n=2)	No (n=11)	Total (n=43)	Yes (n=25)	No (n=18)
Characteristic						
Age at death, mean (SD)	88.36 (3.17)	84.45 (3.03)	89.08 (2.75)	89.56 (3.00)	89.74 (3.06)	89.32 (2.99)
Level of education (%)						
< High school	7.69	0.00	9.09	2.33	0.00	5.56
High school	15.38	0.00	18.18	2.33	0.00	5.56
Bachelor's degree	38.46	0.00	45.45	44.19	56.00	27.78
≥ Master's degree	38.46	100.00	27.27	51.16	44.00	61.11
Presence of *APOE*-ε4 (%)						
No	38.46	100.00	27.27	74.42	76.00	72.22
Yes	61.54	0.00	72.73	25.58	24.00	27.78
Grammatical complexity (%)						
Quartile 1 (low)	46.15	0.00	54.55	23.26	12.00	38.89
Quartile 2	23.08	50.00	18.18	16.28	24.00	5.56
Quartile 3	7.69	0.00	9.09	27.91	32.00	22.22
Quartile 4 (high)	23.08	50.00	18.18	32.56	32.00	33.33
Grammatical complexity (low vs. higher, %)						
Quartile 1 (low)[1]	46.15	0.00	54.55	23.26	12.00	38.89
Quartiles 2-4 (higher)[2]	53.85	100.00	45.45	76.74	88.00	61.11

[1]Low was defined as the lowest quartile.
[2]Higher was defined as the top three quartiles (quartiles two to four).

Abbreviations: ***APOE*-ε4** = Apolipoprotein-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **SD** = Standard deviation

Table 2b. Participant Characteristics Stratified by Grammatical Complexity, CERAD Criteria (n=56)

	Grammatical Complexity					
	Low			Higher		
	Cognitive Resilience (CERAD Criteria)			Cognitive Resilience (CERAD Criteria)		
	Total (n=16)	Yes (n=3)	No (n=13)	Total (n=40)	Yes (n=24)	No (n=16)
Characteristic						
Age at death, mean (SD)	89.11 (3.37)	89.10 (2.50)	89.11 (3.59)	89.36 (2.98)	89.38 (3.44)	89.32 (2.21)
Level of education (%)						
< High school	12.50	0.00	15.38	0.00	0.00	0.00
High school	12.50	0.00	15.38	2.50	0.00	6.25
Bachelor's degree	31.25	66.67	23.08	47.50	50.00	43.75
≥ Master's degree	43.75	33.33	46.15	50.00	50.00	50.00
Presence of *APOE*-ε4 (%)						
No	68.75	100.00	61.54	65.00	75.00	50.00
Yes	31.25	0.00	68.46	35.00	25.00	50.00
Idea density (%)						
Quartile 1 (low)	37.50	0.00	46.15	17.50	8.33	31.25
Quartile 2	18.75	33.33	15.38	22.50	25.00	18.75
Quartile 3	25.00	33.33	23.08	20.00	20.83	18.75
Quartile 4 (high)	18.75	33.33	15.38	40.00	45.83	31.25
Idea density (low vs. higher, %)						
Quartile 1 (low)[1]	37.50	0.00	46.15	17.50	8.33	31.25
Quartiles 2-4 (higher)[2]	62.50	100.00	53.85	82.50	91.67	68.75

[1]Low was defined as the lowest quartile.
[2]Higher was defined as the top three quartiles (quartiles two to four).

Abbreviations: ***APOE*-ε4** = Apolipoprotein-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **SD** = Standard deviation

Table 2c. Participant Characteristics Stratified by Idea Density, NIA-RI Criteria (n=42)

	Idea Density					
	Low			Higher		
	Cognitive Resilience (NIA-RI Criteria)			Cognitive Resilience (NIA-RI Criteria)		
	Total (n=16)	Yes (n=0)	No (n=16)	Total (n=26)	Yes (n=13)	No (n=16)
Characteristic						
Age at death, mean (SD)	88.14 (4.36)	-	88.14 (4.36)	89.84 (2.50)	90.38 (2.61)	89.31 (2.37)
Level of education (%)						
< High school	6.25	-	6.25	3.85	0.00	7.69
High school	12.50	-	12.50	0.00	0.00	0.00
Bachelor's degree	43.75	-	43.75	46.15	61.54	30.77
≥ Master's degree	37.50	-	37.50	50.00	38.46	61.54
Presence of *APOE*-ε4 (%)						
No	43.75	-	43.75	65.38	69.23	61.54
Yes	56.25	-	56.25	34.62	30.77	38.46
Grammatical complexity (%)						
Quartile 1 (low)	56.25	-	56.25	23.08	15.38	30.77
Quartile 2	12.50	-	12.50	15.38	23.08	7.69
Quartile 3	18.75	-	18.75	30.77	30.77	30.77
Quartile 4 (high)	12.50	-	12.50	30.77	30.77	30.77
Grammatical complexity (low vs. higher, %)						
Quartile 1 (low)[1]	56.25	-	56.25	23.08	15.38	30.77
Quartiles 2-4 (higher)[2]	43.75	-	43.75	76.92	84.62	69.23

[1]Low was defined as the lowest quartile.
[2]Higher was defined as the top three quartiles (quartiles two to four).

Abbreviations: ***APOE*-ε4** = Apolipoprotein-ε4; **NIA-RI** = National Institute of Aging – Reagan Institute; **SD** = Standard deviation

Table 2d. Participant Characteristics Stratified by Grammatical complexity, NIA-RI Criteria (n=42)

	Grammatical Complexity					
	Low			Higher		
	Cognitive Resilience (NIA-RI Criteria)			Cognitive Resilience (NIA-RI Criteria)		
	Total (n=15)	Yes (n=2)	No (n=13)	Total (n=27)	Yes (n=11)	No (n=16)
Characteristic						
Age at death, mean (SD)	87.73 (4.47)	90.23 (2.20)	87.34 (4.66)	90.01 (2.32)	90.41 (2.77)	89.73 (2.00)
Level of education (%)						
< High school	13.33	0.00	15.38	0.00	0.00	0.00
High school	13.33	0.00	15.38	0.00	0.00	0.00
Bachelor's degree	33.33	100.00	23.08	51.85	54.55	50.00
≥ Master's degree	40.00	0.00	46.15	48.15	45.45	50.00
Presence of *APOE*-ε4 (%)						
No	60.00	100.00	53.85	55.56	63.64	50.00
Yes	40.00	0.00	46.15	44.44	36.36	50.00
Idea density (%)						
Quartile 1 (low)	60.00	0.00	69.23	25.93	0.00	43.75*
Quartile 2	6.67	0.00	7.69	25.93	45.45	12.50
Quartile 3	20.00	50.00	15.38	11.11	9.09	12.50
Quartile 4 (high)	13.33	50.00	7.69	37.04	45.45	31.25
Idea density (low vs. higher, %)						
Quartile 1 (low)[1]	60.00	0.00	69.23	25.93	0.00	43.75*
Quartiles 2-4 (higher)[2]	40.00	100.00	30.77	74.07	100.00	56.25

* $p<.05$

[1]Low was defined as the lowest quartile.

[2]Higher was defined as the top three quartiles (quartiles two to four).

Note: Overall idea density (quartiles 1-4) and low idea density (lowest quartile versus higher three quartiles) differed significantly by cognitive resilience status among individuals with higher grammatical complexity.

Abbreviations: ***APOE*-ε4** = Apolipoprotein-ε4; **NIA-RI** = National Institute of Aging – Reagan Institute; **SD** = Standard deviation

5.2.2 Research Question 2: Are early-life written language skills associated with cognitive resilience, and does this association hold when adjusted for standard confounders (age and *APOE*)?

Unadjusted logistic regression analyses were used to assess the relationship between written language skills and cognitive resilience in both the CERAD and NIA-RI samples. In the CERAD sample, both low idea density (odds ratio (OR): 0.13, 95% confidence interval (CI): 0.02-0.56) and low grammatical complexity (OR:0.15, 95% CI: 0.03-0.57) were negatively associated with cognitive resilience in later life (Tables 3a and 3b). In other words, individuals with higher idea density had seven times the odds of having cognitive resilience in later life in comparison to those with lower idea density. In Table 3b, those with higher grammatical complexity had six times the odds of having cognitive resilience in comparison to those with low grammatical complexity. Since low idea density and low grammatical complexity were both significant predictors of cognitive resilience using CERAD criteria, the CERAD sample was further explored to determine if the unadjusted association remained after adjustment for standard covariates (i.e., age, *APOE*).

In Tables 3a and 3b, the association of written language skills and cognitive resilience (based on CERAD criteria) was assessed controlling for standard covariates using logistic regression. In the final model of Table 3a, when covariates age and *APOE* were added to both models, both low idea density (OR: 0.15, 95% CI: 0.02-0.72) and low grammatical complexity (OR: 0.13, 95% CI: 0.03-0.50) remained significant. That is, higher idea density was associated with six times the odds of cognitive resilience in comparison to low idea density. With respect to grammatical complexity, the odds of cognitive resilience in later life increased seven-fold among those with higher grammatical complexity compared to those with low grammatical complexity. In Table 3b, the presence of *APOE*-ε4 (OR: 0.28, 95% CI: 0.07-0.95) was also significantly and negatively associated with cognitive resilience in the final model.

In the NIA-RI sample, the association between idea density and cognitive resilience could not be assessed in logistic regression models since no individuals in the lowest quartile of idea density showed cognitive resilience. In the unadjusted model presented in Table 3c, low grammatical complexity (OR: 0.22, 95% CI: 0.03-1.03) was not significantly associated with cognitive resilience in the NIA-RI sample. Similarly, in the full model (adjusted for age and

APOE), low grammatical complexity (OR: 0.25, 95% CI: 0.03-1.25) was not associated with the outcome of interest (based on NIA-RI criteria).

To address some of the limitations of small sample sizes specifically with respect to the NIA-RI sample, analyses using a combined sample of meeting either CERAD or NIA-RI criteria were conducted. Results using the sample of meeting either CERAD or NIA-RI were generally consistent with the results of the CERAD sample and are presented in Appendix D.

Table 3a. The Effect of Age and Apolipoprotein E on the Association of Idea Density and Cognitive Resilience, CERAD Criteria (n=56)

	Cognitive Resilience (CERAD Criteria)			
	Unadjusted Model	Age	*APOE*-ε4 Presence	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable				
Idea Density	**0.13**	**0.13**	**0.16**	**0.15**
Low[1] *vs. high*	**(0.02-0.56)**	**(0.02-0.55)**	**(0.02-0.73)**	**(0.02-0.72)**
Age		0.97		0.97
(years)		(0.80-1.17)		(0.80-1.18)
APOE-ε4			0.51	0.52
Presence			(0.14-1.81)	(0.14-1.83)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **OR** = Odds ratio

Table 3b. The Effect of Age and Apolipoprotein E on the Association of Grammatical Complexity and Cognitive Resilience, CERAD Criteria (n=56)

	Cognitive Resilience (CERAD Criteria)			
	Unadjusted Model	Age	*APOE*-ε4 Presence	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable				
Grammatical Complexity *Low*[1] *vs. high*	**0.15 (0.03-0.57)**	**0.15 (0.03-0.57)**	**0.13 (0.03-0.50)**	**0.13 (0.03-0.50)**
Age (years)		1.01 (0.83-1.22)		1.01 (0.83-1.23)
APOE-ε4 Presence			**0.28 (0.07-0.95)**	**0.28 (0.07-0.95)**

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **OR** = Odds ratio

Table 3c. The Effect of Age and Apolipoprotein E on the Association of Grammatical Complexity and Cognitive Resilience, NIA-RI Criteria (n=42)

	Cognitive Resilience (NIA-RI Criteria)			
	Unadjusted Model	Age	*APOE*-ε4 Presence	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable				
Grammatical Complexity *Low[1] vs. high*	0.22 (0.03-1.03)	0.28 (0.04-1.35)	0.21 (0.03-1.00)	0.25 (0.03-1.25)
Age (years)		1.17 (0.91-1.59)		1.15 (0.88-1.57)
APOE-ε4 Presence			0.42 (0.09-1.71)	0.47 (0.10-1.98)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein-ε4; **CI** = Confidence interval; **NIA-RI** = National Institute of Aging – Reagan Institute; **OR** = Odds ratio

5.2.3 Research Question 3: Does the association between early-life written language skills and later-life cognitive resilience hold when adjusting for indicators of cognitive reserve?

Cognitive Reserve

Cognitive reserve theories suggest that the brain is able to actively cope with neuropathologic damage through compensatory mechanisms such as life experience and education (Stern, 2012). Education was used to reflect cognitive reserve for research question three. Since education could not be assessed using standard logistic regression because of the small sample and a lack of variability in level of education, Firth logistic regression techniques were utilized for Tables 4a and 4b. In Table 4a, idea density was significant in the unadjusted model. When education was added to the model, low idea density (OR: 0.19, 95% CI: 0.04-0.93) remained negatively associated with cognitive resilience. However, education was not significant. In the final model (with covariates age, education, and *APOE*), low idea density (OR: 0.24, 95% CI: 0.04-1.30) was no longer significant.

In the unadjusted model in Table 4b, low grammatical complexity (OR: 0.18, 95% CI: 0.05-0.69) was negatively associated with cognitive resilience. When education was added to the model, grammatical complexity (OR: 0.25, 95% CI: 0.06-0.997) remained significant with a wider confidence interval, although education was not a significant covariate. Grammatical complexity remained significant in adjusted models with covariates age and *APOE*. In the final model, low grammatical complexity (OR: 0.23, 95% CI: 0.05-0.93) remained significantly associated with cognitive resilience. That is, those with higher grammatical complexity had four times greater odds of reflecting cognitive resilience in later life in comparison to those with low grammatical complexity. The presence of *APOE*-ε4 (OR: 0.26, 95% CI: 0.07-0.94) was also significantly and negatively associated with cognitive resilience in the final model.

Table 4a. The Impact of Education on the Association of Idea Density and Cognitive Resilience using Firth Logistic Regression, CERAD Criteria (n=56)

	Cognitive resilience (CERAD Criteria)					
	Unadjusted Model	Age	*APOE*-ε4 Presence	Age and *APOE*-ε4 Presence	Education	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable						
Idea density *Low*[1] *vs. high*	**0.16 (0.03-0.74)**	**0.16 (0.03-0.74)**	**0.20 (0.04-0.95)**	**0.20 (0.04-0.96)**	**0.19 (0.04-0.93)**	0.24 (0.04-1.30)
Age (years)		0.97 (0.81-1.17)		0.98 (0.81-1.18)		0.93 (0.76-1.15)
APOE-ε4 Presence			0.53 (0.15-1.84)	0.54 (0.16-1.86)		0.43 (0.12-1.55)
Education ≤ *High school vs. Master's degree*					0.14 (0.01-4.33)	0.10 (0.003-3.19)
Bachelor's degree vs. Master's degree					1.60 (0.50-5.14)	1.81 (0.53-6.19)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **OR** = Odds ratio

Table 4b. The Impact of Education on the Association of Grammatical Complexity and Cognitive Resilience using Firth Logistic Regression, CERAD Criteria (n=56)

	Cognitive Resilience (CERAD Criteria)					
	Unadjusted Model	Age	*APOE*-ε4 Presence	Age and *APOE*-ε4 Presence	Education	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable						
Grammatical Complexity *Low[1] vs. high*	**0.18 (0.05-0.69)**	**0.18 (0.05-0.70)**	**0.15 (0.04-0.62)**	**0.16 (0.04-0.64)**	**0.25 (0.06-0.997)**	**0.23 (0.05-0.93)**
Age (years)		1.01 (0.84-1.21)		1.01 (0.84-1.22)		0.97 (0.79-1.19)
APOE-ε4 Presence			0.30 (0.09-1.05)	0.30 (0.09-1.06)		**0.26 (0.07-0.94)**
Education *≤ High school vs. Master's degree*					0.18 (0.01-5.30)	0.12 (0.004-4.23)
Bachelor's degree vs. Master's degree					1.42 (0.45-4.50)	1.63 (0.47-5.70)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **OR** = Odds ratio

5.2.3 Research Question 4: Does the association between early-life written language skills and later-life cognitive resilience hold when adjusting for indicators of brain reserve?

Brain Reserve

Brain reserve theories suggest that while instances of brain damage accumulate over time, individuals who show brain reserve may be better able to tolerate AD neuropathology as well as structural changes to the brain (Stern, 2012). Two major structural changes to the brain include atrophy (loss of neuronal tissue, i.e., brain shrinkage), and infarcts, which have been associated with losses in cognitive function as well as AD in later life (Grinberg & Heisen, 2010). To incorporate these theories, infarcts and atrophy were taken into consideration in analyses as indicators of brain reserve. The results of multivariable logistic regression analyses considering brain reserve are presented in Tables 5a through to Table 7d.

Cerebral Infarcts

The presence of cerebral infarcts and their influence on the relationship between written language skills and cognitive resilience was assessed in Tables 5a and 5b. In separate unadjusted models, both low idea density and low grammatical complexity were significantly negatively associated with cognitive resilience in later life. When presence of infarcts was added as a covariate to both models, low idea density (OR: 0.15, 95% CI: 0.02-0.63) and low grammatical complexity (OR: 0.17, 95% CI: 0.03-0.64) remained significant. In separate models presented in Table 5c and 5d, low idea density (OR: 0.14, 95% CI: 0.02-0.63) and low grammatical complexity (OR: 0.17, 95% CI: 0.03-0.63) were also significant when adjusted for number of infarcts. In Tables 5a through to 5d, both presence of infarcts and number of infarcts were not independently significantly associated with cognitive resilience in their separate models or in the final models. However, both low idea density (OR: 0.17, 95% CI: 0.02-0.82 (*presence of infarcts*); OR: 0.16, 95% CI: 0.02-0.77 (*number of infarcts*)) and low grammatical complexity (OR: 0.14, 95% CI: 0.03-0.56 (*presence of infarcts*); OR: 0.14, 95% CI: 0.03-0.55 (*number of infarcts*)) remained significant in their full models (adjusted for age, presence of *APOE*-ε*4,* and presence or number of infarcts) presented in Tables 5a, 5b, 5c and 5d. These final models suggest that individuals with higher idea density had six times the odds of having cognitive resilience in later life in comparison to those with lower idea density. Final models for grammatical complexity also suggest that those with higher grammatical complexity had seven

times the odds of having cognitive resilience in later life in comparison to those with low grammatical complexity. In the full model assessing grammatical complexity and cognitive resilience in Table 5b, *APOE*-ε*4* (OR: 0.28, 95% CI: 0.07-0.97) also became negatively associated with cognitive resilience.

Table 5a. The Association Between Idea Density and Cognitive Resilience Adjusted for Presence of Cerebral Infarcts, CERAD Criteria (n=55)

	Cognitive Resilience (CERAD Criteria)				
	Unadjusted Model	Age	*APOE*-ε4 Presence	Presence of Infarcts	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable					
Idea Density	**0.14**	**0.13**	**0.16**	**0.15**	**0.17**
Low[1] *vs. high*	**(0.02-0.59)**	**(0.02-0.56)**	**(0.02-0.75)**	**(0.02-0.63)**	**(0.02-0.82)**
Age (years)		0.94 (0.76-1.14)			0.98 (0.78-1.20)
APOE-ε4 Presence			0.53 (0.15-1.89)		0.51 (0.14-1.85)
Presence of Infarcts (yes vs. no)				0.52 (0.13-1.92)	0.51 (0.12-2.08)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **OR** = Odds ratio

Table 5b. The Association Between Grammatical Complexity and Cognitive Resilience Adjusted for Presence of Cerebral Infarcts, CERAD Criteria (n=55)

	Cognitive Resilience (CERAD Criteria)				
	Unadjusted Model	Age	*APOE*-ε4 Presence	Presence of Infarcts	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable					
Grammatical Complexity *Low*[1] *vs. high*	**0.16 (0.03-0.59)**	**0.16 (0.03-0.59)**	**0.13 (0.03-0.53)**	**0.17 (0.03-0.64)**	**0.14 (0.03-0.56)**
Age (years)		0.98 (0.79-1.19)			1.03 (0.82-1.28)
APOE-ε4 Presence			**0.29 (0.08-0.96)**		**0.28 (0.07-0.97)**
Presence of Infarcts (yes vs. no)				0.53 (0.13-1.96)	0.46 (0.10-1.95)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **OR** = Odds ratio

Table 5c. The Association Between Idea Density and Cognitive Resilience Adjusted for Number of Cerebral Infarcts, CERAD criteria (n=55)

	Cognitive Resilience (CERAD Criteria)				
	Unadjusted Model	Age	*APOE*-ε4 Presence	Number of Infarcts	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable					
Idea Density	**0.14**	**0.13**	**0.16**	**0.14**	**0.16**
Low[1] *vs. high*	**(0.02-0.59)**	**(0.02-0.56)**	**(0.02-0.75)**	**(0.02-0.63)**	**(0.02-0.77)**
Age (years)		0.94 (0.76-1.14)			0.95 (0.76-1.17)
APOE-ε4 Presence			0.53 (0.15-1.89)		0.54 (0.15-1.95)
Number of Infarcts				0.93 (0.59-1.38)	0.95 (0.60-1.43)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **OR** = Odds ratio

Table 5d. The Association Between Grammatical Complexity and Cognitive Resilience Adjusted for Number of Cerebral Infarcts, CERAD Criteria (n=55)

	Cognitive Resilience (CERAD Criteria)				
	Unadjusted Model	Age	*APOE*-ε4 Presence	Number of Infarcts	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable					
Grammatical Complexity *Low*[1] *vs. high*	**0.16 (0.03-0.59)**	**0.16 (0.03-0.59)**	**0.13 (0.03-0.53)**	**0.17 (0.03-0.63)**	**0.14 (0.03-0.55)**
Age (years)		0.98 (0.79-1.19)			1.00 (0.80-1.23)
APOE-ε4 Presence			0.29 (0.08-1.00)		0.29 (0.08-1.01)
Number of Infarcts				0.93 (0.60-1.36)	0.94 (0.59-1.44)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **OR** = Odds ratio

Cortical Atrophy

When assessing the association between written language skills and cognitive resilience adjusted for cortical atrophy, a smaller sample (n=52) was used due to missing data on cortical atrophy status. Tables 6a and 6b consider the impact of cortical atrophy. In the unadjusted models, both idea density and grammatical complexity were significantly associated with cognitive resilience. In both tables, low idea density (OR: 0.17, 95% CI: 0.02-0.78) and low grammatical complexity (OR: 0.22, 95% CI: 0.04-0.87) were significantly and negatively associated with cognitive resilience when adjusted for cortical atrophy. In the full model in Table 6a, low idea density (OR: 0.18, 95% CI: 0.03-1.02) was not significant when adjusted for age, cortical atrophy and *APOE.* However, in Table 6b, low grammatical complexity (OR: 0.18, 95% CI: 0.03-0.76) remained significantly associated with cognitive resilience in the final model. Presence of *APOE* -ε*4* (OR: 0.26, 95% CI: 0.06-0.97) also was significantly associated with cognitive resilience in the full model.

Table 6a. The Association Between Idea Density and Cognitive Resilience Adjusted for Cortical Atrophy, CERAD Criteria (n=52)

	Cognitive Resilience (CERAD Criteria)				
	Unadjusted Model	Age	*APOE*-ε4 Presence	Cortical Atrophy	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable					
Idea Density	**0.14**	**0.13**	**0.18**	**0.17**	0.18
Low[1] vs. high	**(0.02-0.61)**	**(0.02-0.59)**	**(0.03-0.83)**	**(0.02-0.78)**	(0.03-1.02)
Age (years)		0.95 (0.76-1.16)			0.95 (0.76-1.17)
APOE-ε4 Presence			0.43 (0.11-1.61)		0.45 (0.12-1.76)
Cortical Atrophy *(yes vs. no)*				0.34 (0.09-1.19)	0.33 (0.08-1.17)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **OR** = Odds ratio

Table 6b. The Association Between Grammatical Complexity and Cognitive Resilience Adjusted for Cortical Atrophy, CERAD Criteria (n=52)

	Cognitive Resilience (CERAD Criteria)				
	Unadjusted Model	Age	*APOE*-ε4 Presence	Cortical Atrophy	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable					
Grammatical Complexity *Low[1] vs. high*	**0.16 (0.04-0.72)**	**0.19 (0.04-0.72)**	**0.16 (0.03-0.65)**	**0.22 (0.04-0.87)**	**0.18 (0.03-0.76)**
Age (years)		0.99 (0.80-1.20)			0.98 (0.78-1.22)
APOE-ε4 Presence			**0.25 (0.06-0.90)**		**0.26 (0.06-0.97)**
Cortical Atrophy *(yes vs. no)*				0.32 (0.09-1.10)	0.32 (0.08-1.18)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **OR** = Odds ratio

Infarcts and Atrophy

To assess the impact of both infarcts and atrophy on the relationship between early-life written language skills and cognitive resilience, both covariates were used in the models presented in Tables 7a through to 7d. In Table 7a, low idea density (OR: 0.14, 95% CI: 0.02-0.61) was significantly associated with cognitive resilience in the unadjusted model. When idea density was separately adjusted for the presence of infarcts and cortical atrophy, idea density remained significant. However, in the final model, the association between idea density and cognitive resilience was suggestive but not significant when adjusted for age, presence of infarcts, cortical atrophy, and *APOE*-ε4 (OR: 0.22, 95% CI: 0.03-1.20).

In Table 7b, the presence of infarcts and cortical atrophy was assessed with respect to the association between grammatical complexity and cognitive resilience. In the unadjusted model, low grammatical complexity (OR: 0.19, 95% CI: 0.04-0.72) was significantly associated with cognitive resilience. When presence of infarcts and cortical atrophy were independently added to the model, low grammatical complexity remained negatively associated with cognitive resilience. In the full model (adjusting for age, presence of infarcts, cortical atrophy, and *APOE*-ε4), low grammatical complexity was significantly associated with cognitive resilience in later life (OR: 0.19, 95% CI: 0.03-0.87). *APOE*-ε4 (OR: 0.25, 95% CI: 0.06-0.97) also was significantly associated with showing cognitive resilience in the full model.

In Table 7c and 7d, the association between written language skills (idea density and grammatical complexity) and cognitive resilience incorporating both number of infarcts and cortical atrophy was analyzed. In the separate unadjusted models, both low idea density (OR: 0.14, 95% CI: 0.02-0.61) and low grammatical complexity (OR: 0.19, 95% CI: 0.04-0.72) were significantly and negatively associated with cognitive resilience. When number of infarcts and cortical atrophy were separately added, both idea density and grammatical complexity remained negative but significant predictors of cognitive resilience. In the final model, when age, number of infarcts, cortical atrophy and *APOE* were added, low idea density (OR: 0.21, 95% CI: 0.03-1.08) was no longer significant. However, in the final model in Table 7d, low grammatical complexity (OR: 0.18, 95% CI: 0.03-0.80) was still negatively associated with cognitive resilience in later life. The presence of *APOE*-ε4 (OR: 0.26, 95% CI: 0.06-0.98) was also negatively associated with cognitive resilience in the final model of Table 7d.

A complete model assessing the influence of both cognitive reserve (i.e., education) and

brain reserve (i.e., infarcts and atrophy) could not be completed due to small sample and cell sizes.

Table 7a. The Association Between Idea Density and Cognitive Resilience Adjusted for Presence of Cerebral Infarcts and Cortical Atrophy, CERAD Criteria (n=52)

	Cognitive Resilience (CERAD criteria)					
	Unadjusted Model	Age	*APOE*-ε4 Presence	Presence of Infarcts	Cortical Atrophy	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable						
Idea Density	**0.14**	**0.13**	**0.18**	**0.15**	**0.17**	0.22
Low[1] vs. high	**(0.02-0.61)**	**(0.02-0.59)**	**(0.03-0.83)**	**(0.02-0.67)**	**(0.02-0.78)**	(0.03-1.20)
Age (years)		0.95 (0.77-1.16)				1.00 (0.79-1.24)
APOE-ε4 Presence			0.43 (0.11-1.61)			0.41 (0.09-1.61)
Presence of Infarcts *(yes vs. no)*				0.41 (0.09-1.60)		0.32 (0.06-1.47)
Cortical Atrophy *(yes vs. no)*					0.34 (0.09-1.19)	0.30 (0.07-1.09)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **OR** = Odds ratio

Table 7b. The Association Between Grammatical Complexity and Cognitive Resilience Adjusted for Presence of Cerebral Infarcts and Cortical Atrophy, CERAD Criteria (n=52)

	Cognitive Resilience (CERAD Criteria)					
	Unadjusted Model	Age	*APOE*-ε4 Presence	Presence of Infarcts	Cortical Atrophy	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable						
Grammatical Complexity *Low[1] vs. high*	**0.19 (0.04-0.72)**	**0.19 (0.04-0.72)**	**0.16 (0.03-0.65)**	**0.21 (0.04-0.82)**	**0.22 (0.04-0.87)**	**0.19 (0.03-0.87)**
Age (years)		0.99 (0.80-1.20)				1.03 (0.81-1.32)
APOE-ε4 Presence			**0.25 (0.06-0.90)**			**0.25 (0.06-0.97)**
Presence of Infarcts *(yes vs. no)*				0.43 (0.10-1.69)		0.33 (0.06-1.57)
Cortical Atrophy *(yes vs. no)*					0.32 (0.09-1.10)	0.30 (0.07-1.13)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **OR** = Odds ratio

Table 7c. The Association Between Idea Density and Cognitive Resilience Adjusted for Number of Cerebral Infarcts and Cortical Atrophy, CERAD Criteria (n=52)

	Cognitive Resilience (CERAD Criteria)					
	Unadjusted Model	Age	*APOE*-ε4 Presence	Number of Infarcts	Cortical Atrophy	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable						
Idea Density *Low[1] vs. high*	**0.14 (0.02-0.61)**	**0.13 (0.02-0.59)**	**0.18 (0.03-0.83)**	**0.15 (0.02-0.66)**	**0.17 (0.02-0.78)**	0.21 (0.03-1.08)
Age (years)		0.95 (0.77-1.16)				0.95 (0.76-1.19)
APOE-ε4 Presence			0.43 (0.11-1.61)			0.45 (0.11-1.75)
Number of Infarcts				0.90 (0.56-1.34)		0.94 (0.56-1.44)
Cortical Atrophy *(yes vs. no)*					0.34 (0.09-1.19)	0.33 (0.08-1.20)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **OR** = Odds ratio

Table 7d. The Association Between Grammatical Complexity and Cognitive Resilience Adjusted for Number of Cerebral Infarcts and Cortical Atrophy, CERAD Criteria (n=52)

	Cognitive Resilience (CERAD Criteria)					
	Unadjusted Model	Age	*APOE*-ε4 Presence	Number of Infarcts	Cortical Atrophy	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable						
Grammatical Complexity *Low[1] vs. high*	**0.19 (0.04-0.72)**	**0.19 (0.04-0.72)**	**0.16 (0.03-0.65)**	**0.20 (0.04-0.78)**	**0.22 (0.04-0.87)**	**0.18 (0.03-0.80)**
Age (years)		0.99 (0.80-1.20)				0.99 (0.78-1.24)
APOE-ε4 Presence			**0.25 (0.06-0.90)**			**0.26 (0.06-0.98)**
Number of Infarcts				0.90 (0.57-1.33)		0.95 (0.58-1.47)
Cortical Atrophy *(yes vs.no)*					0.32 (0.09-1.10)	0.33 (0.08-1.21)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **OR** = Odds ratio

Logistic regression models of the relationship between early-life written language skills and cognitive resilience suggested a consistent association between grammatical complexity and cognitive resilience. In the CERAD sample, grammatical complexity was consistently associated with cognitive resilience in the presence of standard covariates as well as education, cortical atrophy and cerebral infarcts. However, findings on the association between idea density and cognitive resilience were mixed. In the CERAD sample, idea density was significantly associated with cognitive resilience in the presence of standard covariates. However, when adjusted for additional factors such as education, presence of infarcts or cortical atrophy, idea density was no longer associated with cognitive resilience. In Appendix D, the association between written language skills and cognitive resilience was assessed using a larger sample meeting either CERAD or NIA-RI neuropathologic criteria. Using this larger sample, idea density was consistently associated with cognitive resilience even in the presence of covariates. A summary of the analyses for the two samples (CERAD sample, and sample meeting either CERAD or NIA-RI criteria) for research questions two, three and four is provided in Table 8.

Table 8. Summary of Findings on the Association Between Early-life Written Language Skills and Cognitive Resilience

Question #	Results Tables	Exposure	Covariates	CERAD Criteria (n=56)	Either CERAD or NIA-RI Criteria (n=62)
2	Table 3a	Low idea density	Unadjusted	↓	↓
	Table D1	Low idea density	Adjusted by age	↓	↓
		Low idea density	Adjusted by *APOE*	↓	↓
		Low idea density	Adjusted by age and *APOE*	↓	↓
2	Table 3b	Low grammatical complexity	Unadjusted	↓	↓
	Table D2	Low grammatical complexity	Adjusted by age	↓	↓
		Low grammatical complexity	Adjusted by *APOE*	↓	↓
		Low grammatical complexity	Adjusted by age and *APOE*	↓	↓
3	Table 4a	Low idea density	Unadjusted	↓	↓
	Table D3	Low idea density	Adjusted by age	↓	↓
		Low idea density	Adjusted by *APOE*	↓	↓
		Low idea density	Adjusted by age and *APOE*	↓	↓
		Low idea density	Adjusted by education	↓	↓
		Low idea density	Adjusted by age, *APOE* and education	x	↓
3	Table 4b	Low grammatical complexity	Unadjusted	↓	↓

Question #	Results Tables	Exposure	Covariates	CERAD Criteria (n=56)	Either CERAD or NIA-RI Criteria (n=62)
	Table D4	Low grammatical complexity	Adjusted by age	↓	↓
		Low grammatical complexity	Adjusted by *APOE*	↓	↓
		Low grammatical complexity	Adjusted by age and *APOE*	↓	↓
		Low grammatical complexity	Adjusted by education	↓	↓
		Low grammatical complexity	Adjusted by age, *APOE* and education	↓	↓
4	Table 5a	Low idea density	Unadjusted	↓	↓
	Table D5	Low idea density	Adjusted by age	↓	↓
		Low idea density	Adjusted by *APOE*	↓	↓
		Low idea density	Adjusted by presence of infarcts	↓	↓
		Low idea density	Adjusted by age, *APOE*, and presence of infarcts	↓	↓
4	Table 5b	Low grammatical complexity	Unadjusted	↓	↓
	Table D6	Low grammatical complexity	Adjusted by age	↓	↓
		Low grammatical complexity	Adjusted by *APOE*	↓	↓
		Low grammatical complexity	Adjusted by presence of infarcts	↓	↓

Question #	Results Tables	Exposure	Covariates	CERAD Criteria (n=56)	Either CERAD or NIA-RI Criteria (n=62)
		Low grammatical complexity	Adjusted by age, *APOE*, and presence of infarcts	↓	↓
4	Table 5c	Low idea density	Unadjusted	↓	↓
	Table D7	Low idea density	Adjusted by age	↓	↓
		Low idea density	Adjusted by *APOE*	↓	↓
		Low idea density	Adjusted by number of infarcts	↓	↓
		Low idea density	Adjusted by age, *APOE*, and number of infarcts	↓	↓
4	Table 5d	Low grammatical complexity	Unadjusted	↓	↓
	Table D8	Low grammatical complexity	Adjusted by age	↓	↓
		Low grammatical complexity	Adjusted by *APOE*	↓	↓
		Low grammatical complexity	Adjusted by number of infarcts	↓	↓
		Low grammatical complexity	Adjusted by age, *APOE*, and number of infarcts	↓	↓
4	Table 6a	Low idea density	Unadjusted	↓	↓
	Table D9	Low idea density	Adjusted by age	↓	↓
		Low idea density	Adjusted by *APOE*	↓	↓
		Low idea density	Adjusted by cortical atrophy	↓	↓
		Low idea density	Adjusted by age, *APOE*, cortical atrophy	x	↓

Question #	Results Tables	Exposure	Covariates	CERAD Criteria (n=56)	Either CERAD or NIA-RI Criteria (n=62)
4	Table 6b	Low grammatical complexity	Unadjusted	↓	↓
	Table D10	Low grammatical complexity	Adjusted by age	↓	↓
		Low grammatical complexity	Adjusted by *APOE*	↓	↓
		Low grammatical complexity	Adjusted by cortical atrophy	↓	↓
		Low grammatical complexity	Adjusted by age, *APOE*, cortical atrophy	↓	↓
4	Table 7a	Low idea density	Unadjusted	↓	↓
	Table D11	Low idea density	Adjusted by age	↓	↓
		Low idea density	Adjusted by *APOE*	↓	↓
		Low idea density	Adjusted by presence of infarcts	↓	↓
		Low idea density	Adjusted by cortical atrophy	↓	↓
		Low idea density	Adjusted by age, *APOE*, presence of infarcts and cortical atrophy	x	↓
4	Table 7b	Low grammatical complexity	Unadjusted	↓	↓
	Table D12	Low grammatical complexity	Adjusted by age	↓	↓

Question #	Results Tables	Exposure	Covariates	CERAD Criteria (n=56)	Either CERAD or NIA-RI Criteria (n=62)
		Low grammatical complexity	Adjusted by *APOE*	↓	↓
		Low grammatical complexity	Adjusted by presence of infarcts	↓	↓
		Low grammatical complexity	Adjusted by cortical atrophy	↓	↓
		Low grammatical complexity	Adjusted by age, *APOE*, presence of infarcts and cortical atrophy	↓	↓
4	Table 7c	Low idea density	Unadjusted	↓	↓
	Table D13	Low idea density	Adjusted by age	↓	↓
		Low idea density	Adjusted by *APOE*	↓	↓
		Low idea density	Adjusted by number of infarcts	↓	↓
		Low idea density	Adjusted by cortical atrophy	↓	↓
		Low idea density	Adjusted by age, *APOE*, number of infarcts and cortical atrophy	x	↓
4	Table 7d	Low grammatical complexity	Unadjusted	↓	↓
	Table D13	Low grammatical complexity	Adjusted by age	↓	↓
		Low grammatical complexity	Adjusted by *APOE*	↓	↓

Question #	Results Tables	Exposure	Covariates	CERAD Criteria (n=56)	Either CERAD or NIA-RI Criteria (n=62)
		Low grammatical complexity	Adjusted by number of infarcts	↓	↓
		Low grammatical complexity	Adjusted by cortical atrophy	↓	↓
		Low grammatical complexity	Adjusted by age, *APOE*, number of infarcts and cortical atrophy	↓	↓

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease.

Note: Arrows indicate the direction of the significant association (e.g., upward arrows indicate a positive association, downward arrows indicate a negative association); "x" indicates a non-significant finding.

Chapter 6
Discussion

This study aimed to investigate whether cognitive resilience varied by level of early-life written language skills, and if an association existed between early-life written language skills and cognitive resilience. Previous literature using Nun Study data had examined relationships between early-life written language skills with respect to other cognitive outcomes, such as MMSE scores, dementia or AD (Mitzner & Kemper, 2003, Riley et al., 2005). This research was also conducted on a smaller sample which limited the ability of researchers at the time to assess other cognitive outcomes such as cognitive resilience. Other studies have assessed the relationship between measures of verbal ability or multilingualism with cognitive states such as the amnestic form of mild cognitive impairment. Early hypotheses by Snowdon et al. (1996) suggested that higher levels of linguistic ability could be a potential indicator of cognitive resilience. However, only one individual from their sample met the neuropathologic criteria for AD and did not show the expected decline (i.e., remained cognitively resilient) (Snowdon et al., 1996). Due to additional years of follow-up in the Nun Study, the sample available expanded, and cognitive resilience could be assessed in this study. The potential relationship between written language skills and later-life cognitive resilience had not been previously analyzed. Further, the potential relationship had not been assessed with respect to indicators of cognitive or brain reserve.

There were four major research questions that guided this research. These questions were: 1) to characterize how cognitive resilience varied by level of early-life written language skills, 2) to determine if written language skills were associated with cognitive resilience, and whether the association held when adjusting for standard covariates, and, 3) to determine if the association between early-life written language skills and later-life cognitive resilience held when adjusting for indicators of cognitive reserve and 4) to assess if the association held when adjusting for indicators of brain reserve. Based on previous literature, it was hypothesized that individuals with lower levels of written language skills would be less likely to be cognitively resilient in later life. It was also hypothesized that the association between weak written language skills and cognitive resilience would vary by levels of indicators of cognitive reserve (education) and brain reserve (presence and number of infarcts, cortical atrophy).

6.1 Summary of Findings

This study aimed to examine the four major research questions to better understand the relationship between written language skills and cognitive resilience. In the initial descriptive tables (1a and 1b), chi-square tests determined significant associations between idea density and cognitive resilience in the CERAD sample. Grammatical complexity also had a strong association ($p<.01$) with cognitive resilience in the CERAD sample. In the NIA-RI descriptive Table 1a, idea density was also significantly associated with cognitive resilience ($p<.01$).

Research question one assessed how cognitive resilience varied by level of early-life written language skills. Descriptive analyses suggested that if an individual had low idea density or low grammatical complexity it was unlikely for them to show cognitive resilience in later life. For example, in the CERAD sample, only 15% of individuals with low idea density showed cognitive resilience in later life. Individuals who showed cognitive resilience also tended to have higher education (i.e., a Bachelor's degree or higher) in both the lower idea density and grammatical complexity groups. In the NIA-RI sample, no one had cognitive resilience in the low idea density group. Since no one with low idea density showed cognitive resilience, it suggests that low idea density could in fact be a predictor of cognitive resilience in that higher idea density was necessary in order to show cognitive resilience. For individuals with higher grammatical complexity in the NIA-RI sample, idea density was a significant predictor of cognitive resilience.

In research question two, whether early-life written language skills were associated with cognitive resilience, and whether this association held when adjusted for standard covariates were analyzed. Based on multivariable logistic regression techniques, both idea density and grammatical complexity were significantly associated with cognitive resilience in the CERAD sample. The significant associations determined in the CERAD sample suggest that early-life written language skills were indeed associated with cognitive resilience in later life. However, in the NIA-RI sample, idea density could not be fully assessed because in the low idea density category, no individuals had cognitive resilience. However, in earlier analyses presented in the descriptive tables (1a and 1b), idea density ($p<.01$) was a strong and significant predictor of cognitive resilience in the NIA-RI sample. In the NIA-RI sample, no individuals with low idea density showed cognitive resilience. A lack of individuals with both low idea density and cognitive resilience could suggest that low idea density was a strong predictor of cognitive

resilience. That is, if an individual falls in the lowest quartile of idea density, they do not have any chance of showing cognitive resilience based on NIA-RI criteria. However, results may have been influenced by the small sample size, and an association between idea density and cognitive resilience could not be determined since logistic regression techniques were not possible. In the NIA-RI sample grammatical complexity was not a significant predictor of cognitive resilience in unadjusted or models adjusted for age and *APOE*. To clarify the relationship between written language skills and cognitive resilience, future research should aim to study a larger sample that allows for an assessment of both NPs and NFTs (i.e., utilizing a larger NIA-RI sample for analyses).

To further address research question two, the next step was to assess whether the association between written language skills and cognitive resilience held when adjusted for standard confounders (age and *APOE*). Since grammatical complexity was not significant in the NIA-RI sample, and idea density could not be assessed, further analyses solely focused on the CERAD sample. Analyses suggested that both types of written language skills (idea density and grammatical complexity) were generally associated with cognitive resilience when adjusted for standard confounders. In Tables 3a and 3b, when incorporating age and *APOE*, both low idea density and low grammatical complexity were significant negative predictors of cognitive resilience in both unadjusted and final models.

Since education could not be assessed using standard logistic regression, Firth regression techniques were used to assess the relationship between written language skills and cognitive resilience with education as an indicator of cognitive reserve (Tables 4a and 4b) for research question three. Table 4a revealed that idea density was significant in the unadjusted model, as well as when separate covariates age, *APOE* and education were added. Low idea density was also negatively associated with cognitive resilience when adjusted for both age and *APOE* in the same model. However, in the full model (adjusted for age, education and *APOE* simultaneously) idea density was not significant. The results summarized in Table 4a suggest that idea density is a complex early-life factor and is influenced by multiple covariates. However, using a larger sample of those who met either CERAD or NIA-RI criteria, analyses showed that idea density was significantly associated with cognitive resilience (Appendix D, Table D3). Differences in findings between the association between idea density and cognitive resilience in the presence of standard covariates and education suggest that sample size may have influenced the results. In

comparison, in Table 4b, grammatical complexity was negatively associated with cognitive resilience in the unadjusted and adjusted models including the full model (adjusted for age, *APOE* and education*).* Table 4b suggested that grammatical complexity was less influenced by a combination of covariates in comparison to idea density. Similarly, it was evident from Tables 4a and 4b that idea density was much more influenced by the addition of education as a covariate in comparison to grammatical complexity. The finding that education had a stronger influence on idea density than on grammatical complexity was consistent with previous findings. In 1990, Kemper et al. identified the association between education and idea density in adult narratives. These findings were further supported by Cheung and Kemper (1992) who found that individuals with higher levels of educational attainment were better able to pack ideas into fewer words. Findings from early literature are echoed in Tables 4a and 4b, where the stronger influence of education on the relationship between idea density and cognitive resilience in comparison to grammatical complexity is evident.

Indicators of brain reserve were analyzed in research question four. In the separate adjusted models for presence of infarcts and number of infarcts, both idea density and grammatical complexity remained significant in their separate models. In the final models (adjusting for age, presence of infarcts, number of infarcts and *APOE-ε4*), both low idea density and grammatical complexity were still negatively associated with cognitive resilience in later life.

In Table 6a (assessing cortical atrophy and the association between idea density and cognitive resilience), idea density was significant when only adjusted for cortical atrophy. However, idea density was not significant in the final model (including age, cortical atrophy and *APOE-ε4*). These findings again suggested that idea density was strongly influenced by the multiple factors or covariates. That is, when an individual had a combination of negative characteristics (low idea density, older age, cortical atrophy and *APOE-ε4*), they were very unlikely to show cognitive resilience. However, idea density remained consistently and significantly associated with cognitive resilience (while adjusting for standard covariates and cortical atrophy) when using a larger sample including either CERAD or NIA-RI criteria in Appendix D, suggesting that lack of significance in some models based on the smaller CERAD sample may be due to small sample sizes. In comparison, grammatical complexity was significant when only adjusted for cortical atrophy, as well as in the final model (adjusted for

age, cortical atrophy and *APOE*). *APOE-ε4* was also a significant and negative predictor of cognitive resilience in its separate model, and in the final full model. In Tables 7a through 7d, additional models incorporated both infarcts (presence and number separately) and cortical atrophy. In Table 7a, presence of infarcts and cortical atrophy were considered with respect to the relationship between idea density and cognitive resilience. Idea density was significant in separate models when adjusted for age, presence of infarcts, cortical atrophy and *APOE*. However, in the full model, idea density was not significant when adjusted for all of the covariates. The pattern described in Tables 4a and 4b was also repeated in Table 7c, where idea density was significant in separate models for age, number of infarcts, cortical atrophy and *APOE*. However, idea density was not significant in the fully adjusted final model, which was likely related to the small sample. In comparison, idea density remained consistently associated with cognitive resilience in the larger sample presented in Appendix D (sample meeting either CERAD or NIA-RI criteria).

Findings that idea density was not significant in fully adjusted models for age, *APOE*, number or presence of cerebral infarcts and cortical atrophy, however, could reflect the findings of previous studies assessing changes in the brain and vocabulary or content. Mummery et al. (2000) found that individuals who presented with structural brain changes typically had reduced ability to utilize expressive and receptive vocabulary. Difficulty in "remembering" semantic content (such as people, places, or things) was also associated with cortical atrophy in later life (Mummery et al., 2000). Similar findings suggest that both cerebral infarcts and cortical atrophy were significantly associated with weaker performance on tests of vocabulary (Mummery et al., 2000; Saykin et al., 2006;). In this study, the association between idea density and cognitive resilience was also modified by the presence of structural brain changes. However, findings in Tables 7a and 7c also highlight that a combination of complex factors throughout the life course impact the relationship between idea density and cognitive resilience.

In comparison, in Table 7b and 7d, grammatical complexity was consistently significant in unadjusted as well as separate adjusted models (age, presence or number of infarcts, cortical atrophy and *APOE*). In both tables, low grammatical complexity was also a significant negative predictor of cognitive resilience in full models. These tables further indicate that the relationship between grammatical complexity and cognitive resilience was less impacted by covariates than idea density. These findings could relate to previous literature that suggested an association

between grammatical complexity and working memory and executive function (Cheung & Kemper, 1992). Executive function is a general system that regulates attention and control throughout life (Carpenter et al., 2000). Since the present study suggests a consistent association between low grammatical complexity and being less likely to show cognitive resilience in later life, it is plausible that grammatical complexity is a more stable indicator of cognitive resilience and has a stronger association with cognitive resilience. In comparison, idea density was less consistently associated with cognitive resilience and was influenced by complex interactions between a variety of factors across the life course.

Differences in findings between the associations of idea density and grammatical complexity with cognitive resilience may reflect differences in what each linguistic measure represents. Idea density is a measure of ideas expressed, general knowledge and vocabulary (Kemper et al., 2001b; Snowdon et al., 2000). That is, idea density relates to the content of writing. In contrast, grammatical complexity represents the ability to utilize syntactic elements and therefore represents the structure of writing (Riley et al., 2005; Kemper et al., 2001b). Differences in these measures (i.e., content versus structure) provide clues as to why grammatical complexity was a consistent predictor of cognitive resilience whereas idea density was not. In 2001, Kemper et al. found that (semantic) content and an individual's vocabulary (measures of idea density) tended to decline more rapidly with age. In comparison, grammatical complexity declined at a slower rate (Kemper et al., 2001b). These findings suggest that older adults with advanced dementia were still able to produce grammatically complex sentences even when they had challenges associated with dementia (e.g., word finding and memory problems) in producing content (Kemper et al., 2001b). Findings from Kemper et al. (2001b) are amplified in the present study, which highlights that even with respect to later life, the effect of grammatical complexity remains consistent over time and is a strong predictor of cognitive resilience. In contrast, the association between idea density and cognitive resilience varied in the presence of covariates, and thus could be explained away by these different characteristics.

Sensitivity analyses using a larger sample that included either CERAD or NIA-RI criteria (n=62) were also presented in Appendix D. Sensitivity analyses incorporated standard confounders (i.e., age, *APOE*) and indicators of both cognitive (i.e., education) and brain reserve (i.e., infarcts and atrophy). When comparing the CERAD sample to the analyses presented in Appendix D, some of the sensitivity analyses varied. In particular, inconsistencies were noticed

with respect to the association between idea density and cognitive resilience. In the CERAD sample, low idea density was not significantly associated with cognitive resilience in the presence of standard covariates, education, presence or number of infarcts and cortical atrophy. However, in the sensitivity analyses, low idea density remained consistently and significantly associated with cognitive resilience. Differences in findings could also suggest that the CERAD analytic sample was too small to suggest significant associations.

Although the association between written language skills and cognitive resilience has not been previously assessed, the overall findings were somewhat consistent with prior research. Mitzner & Kemper (2003) determined that idea density was associated with cognitive status as measured by MMSE scores. Similar findings by Mueller et al. (2015) also determined that additional measures from written language samples (e.g., verbal fluency and switching) were associated with cognitive states such as MCI in later life. Earlier work also determined that individuals with low idea density had higher odds of moderate or severe cerebral atrophy as well as meeting the neuropathologic criteria for AD (Riley et al., 2005; Snowdon et al., 1996). Although variables like atrophy were not assessed in the same manner as previous findings, the present study did suggest an association between written language skills (idea density and grammatical complexity) and the long-term cognitive outcome of cognitive resilience. A summary chart of findings is presented in the Results section, Table 8.

6.2 Limitations

Although there were many strengths associated with the present research, there were also limitations. One of the major limitations of using Nun Study data is generalizability. The Nun Study population only consists of females, meaning findings cannot be generalized to males. The sample has a relatively homogenous lifestyle and may therefore differ from the general public in terms of factors such as tobacco use, alcohol consumption, reproductive history, lifestyle and marital status. These factors may thus decrease the generalizability of study findings.

Another limitation may have existed in the assessment of AD using CERAD neuropathologic criteria. As previously discussed, CERAD is a standardized protocol used to assess NPs. Therefore, outcome measures using CERAD may be related to resilience to impacts of NPs rather than to Alzheimer neuropathology overall.

Participation in the Nun Study was restricted to individuals who were 75 years of age or

older at baseline. Thus, individuals who died before then were not included in the study, and this may have led to a survivor bias. This would include exclusion of those who developed AD at an earlier age (e.g., were more susceptible to AD or had early-onset AD).

In addition, the sample was relatively small. In the Nun Study, data on idea density and grammatical complexity rely upon the availability of written autobiographies. Only 180 coded autobiographies were available for further analysis. The presence of cognitive resilience relies upon the presence of Alzheimer neuropathology (as verified by CERAD or NIA-RI), and the absence of the clinical symptoms of dementia. Individuals who meet these criteria were relatively rare. Limitations in sample size also hindered the ability to properly assess written language skills with respect to cognitive resilience for those with low idea density in the NIA-RI sample. Since no individuals with low idea density had cognitive resilience in the NIA-RI sample, the potential relationship could not be determined with the multivariable models. However, analyses using a combined sample of meeting either CERAD or NIA-RI suggested no major differences in terms of the relationship between early-life written language skills and later-life cognitive resilience (Appendix D).

Another limitation was the differences across analytic samples. In Appendix E, comparisons between samples and the Nun Study population are provided. In Table E1, it is evident that age at death and education are significantly different in the total autobiography sample (n=180) in comparison to the total Nun Study population (n=678). The total Nun Study population were significantly older at death and were, in general, more highly educated than the total autobiography sample. In both the CERAD and NIA-RI samples, excluded participants were significantly younger than the individuals in the analytic samples. These findings are not surprising given that participants who were generally younger would have been less likely to have developed AD neuropathology required for inclusion in the analytic sample.

The outcome variable of cognitive resilience also had limitations. Stern (2002) described the concept as being dynamic, and encompassing a broad spectrum between absence and presence. However, the operationalization of cognitive resilience in the present study was binary (present vs. absent). In addition, the criteria for the presence of cognitive resilience only required the absence of dementia rather than a more stringent classification of intact cognition.

6.3 Strengths

The present research provided a variety of strengths that support the ability to assess the relationship of interest. In terms of lifestyle, participants in the Nun Study were largely free of confounding variables, such as heavy alcohol consumption, tobacco use, household income and access to health care resources. For example, Norton et al. (2014) suggest that globally 13.9% of all cases of AD are attributable to smoking. However, given the lifestyle of Nun Study participants (non-smokers), variation and thus potential confounding by tobacco use does not influence the findings of this research.

Another strength was the use of gold-standard neuropathologic criteria for assessing AD. While other studies rely predominantly upon cognitive testing measures or assessment of NPs (through CERAD), neuropathology in the Nun Study accounts for both NPs and NFTs (using CERAD and NIA-RI).

The Nun Study also provides strong data for assessing cognitive resilience. In addition to neuropathologic analyses, the Nun Study has data on the dementia status of participants at last cognitive assessment before death. Having both AD neuropathologic data and cognitive assessments allows operationalization of cognitive resilience. As well, the Nun Study provided useful variables (i.e., education, cortical atrophy, cortical infarcts) for assessing both cognitive and brain reserve.

Lastly, the study design of the Nun Study was a strength as it is a longitudinal, population-based cohort study. In population-based cohort studies, a set population is assessed over a longitudinal period to better understand the relationship between exposures and outcomes (Szklo, 1998). Nun Study data provide information on early-life exposures such as written language skills in addition to annual cognitive assessments.

6.4 Implications and Future Directions

This research aimed to enhance current knowledge on the potential association between early-life written language skills and later-life cognitive resilience. Previous studies suggested a strong relationship between low levels of written language and the occurrence of AD in later life (Mitzner & Kemper, 2003; Mueller et al., 2015). Similarly, written language skills were protective against cognitive impairment (Snowdon et al., 1996). However, limitations in previous research, such as neuropathologic criteria that do not assess both NFTs and NPs, limited the

strength of these findings (Snowdon et al., 1996). Furthermore, limited evidence has assessed written language skills with respect to cognitive resilience. This study provided understanding of the link between written language skills and the development of resilience. Specifically, a consistent association between grammatical complexity and late-life cognitive resilience was determined.

Findings from this research will have a number of practical implications. Since previous research has not assessed the impact of early-life written language skills on resilience, the present research will provide groundwork for future studies. Future studies should aim to assess the content of other forms of writing, such as writing presented in social media with respect to later-life cognitive outcomes. Future generations may be less likely to write autobiographical essays and may prefer to use other types of writing (e.g., social media). In a previous study by Agichtein et al. (2008), the quality of content posted in social media was assessed (i.e., Yahoo! Answers). However, this previous research on social media content could also be expanded to relate these measures of semantic complexity and grammaticality to cognitive resilience in later life. Further research could aim to understand whether other forms of writing (i.e., other than autobiographies) are linked to cognitive outcomes such as cognitive resilience.

Due to limitations in sample size in the present study, future studies should also aim to replicate this study's approach using a larger sample. In future studies, written language skills could also be assessed for their association with structural changes in the brain, memory impairments, or changes in cognitive states (e.g., transitions from MCI to normal or intact cognition). Future studies should also aim to utilize a broader definition of cognitive resilience. For example, future studies could aim to utilize more stringent criteria in terms of defining cognitive resilience as the presence or intact cognition in conjunction with Alzheimer neuropathology.

Other contributions through this research were also made to understanding cognitive resilience. Cognitive resilience is a complex outcome. Therefore, further developments in our understanding of written language skills and their interaction with other variables throughout the life course is beneficial and critical to understanding how individuals develop cognitive resilience.

Since an association was found between written language skills and cognitive resilience, these results will be valuable in guiding policies and interventions to prevent cognitive

impairment. Findings will be helpful in prioritizing interventions in early life, such as improving cognitive stimulation, verbal ability and working memory in schools. This study also provides support for the importance of early development in terms of reading, writing and language acquisition. Additionally, continued support is needed with respect to beneficial early-life circumstances and later-life cognition. Since written autobiographies were predictive of cognitive resilience in later life, these findings provide support for the need for continued focus on language skills and in particular grammatical complexity throughout the life course. Overall, improving written language skills in early life could be a modifiable technique used to develop cognitive resilience and thereby reduce the impact of AD in later life.

References

Abbott, R., Berninger, V. W., & Fayol, M. (2010). Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology*, 102, 281-298.

Agichtein, E., Castillo, C., Donato, D., Gionis, A., & Mishne, G. (2008). Finding high-quality content in social media. *Proceedings of the international conference on web search and data mining*, 183-194, New York, NY, USA.

Alzheimer's Association. (2016). Diagnostic procedure. Retrieved from: http://www.alz.org/professionals_and_researchers_diagnostic_procedures.asp

Alzheimer's Association (2017). Major milestones in Alzheimer's and brain research. Retrieved from: http://www.alz.org/research/science/major_milestones_in_alzheimers.asp

Alzheimer Society of Canada. (2010). Rising tide: the impact of dementia on Canadian society. Retrieved from: http://www.alzheimer.ca/~/media/Files/national/Advocacy/ASC_Rising_Tide_Full_Report_e.pdf

Alzheimer Society of Canada. (2016). Executive summary – prevalence and monetary costs of dementia in Canada. Retrieved from: http://www.alzheimer.ca/~/media/Files/national/Statistics/prevalence_summary_e.pdf

Alzheimer Society of Canada. (2017). About dementia. Retrieved from: http://www.alzheimer.ca/en/About-dementia

Allen, R. S., Haley, P. P., Harris, G. M., Fowler, S. N., & Pruthi, R. (2011). Resilience: definitions, ambiguities, and applications. In *Resilience in Aging* (pp. 1-13). New York: Springer.

Anstey, K. J., Mack, H. A., & Cherbuin, N. (2009). Alcohol consumption as a risk factor for dementia and cognitive decline: meta-analysis of prospective studies. *American Journal of Geriatric Psychiatry*, *17*(7), 542-555.

Arif, S., & Albulene, G. (2016). The relation between socio-economic status (SES) and early development: empirical findings and theoretical perspectives. *Socioeconomica,* 5(10), 309-329.

Arriaga, R. I., Fenson, L., Cronan, T., & Pethick, S. J. (1998). Scores on the MacArthur communicative development inventory of children from low- and middle-income families. *Applied Psycholinguistics*, *19*(02), 209-223.

Baumgart, M., Snyder, H. M., Carrillo, M. C., Fazio, S., Kim, H., & Johns, H. (2015). Summary of the evidence on modifiable risk factors for cognitive decline and dementia: A population-based perspective. *Alzheimer's & Dementia*, *11*(6), 718-726.

Berninger, V., Abbott, R., Swanson, H. L., Lovitt, D., Trivedi, P., Lin, S., Trivedi, P., Lin, S., Gould, L., Youngstrom, M., Shimada, S., & Amtmann, D. (2010). Relationship of word- and sentence-level working memory to reading and writing in second, fourth, and sixth grade. *American Speech-Language Hearing Association*, 41, 179-193.

Bialystok, E., Craik, F. I., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia, 45*(2), 459-464.

Braak, H., & Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathologica*, *82*(4), 239-259.

Brewster, P. W., Melrose, R. J., Marquine, M. J., Johnson, J. K., Napoles, A., MacKay-Brandt, A., Farias, S., Reed, B., & Mungas, D. (2014). Life experience and demographic influences on cognitive function in older adults. *Neuropsychology*, *28*(6), 846-858.

Brown, C., Snodgrass, T., Kemper, S. J., Herman, R., & Covington, M. A. (2008). Automatic measurement of propositional idea density from part-of-speech tagging. *Behavior Research Methods*, *40*(2), 540-545.

Canadian Institute for Health Information. (2011). Health care in Canada, 2011 a focus on seniors and aging. Retrieved from: https://secure.cihi.ca/free_products/HCIC_2011_seniors_report_en.pdf

Carpenter, P. A., Just, M. A., & Reichle, E. D. (2000). Working memory and executive function: evidence from neuroimaging. *Current Opinion in Neurobiology*, *10*(2), 195-199.

Carrillo, M. C., Thies, W., & Bain, L. J. (2012). The global impact of Alzheimer's disease. In *Alzheimer's Disease-Modernizing Concept, Biological Diagnosis and Therapy*, *28,* 1-14. Frankfurt: Karger Publishers.

Chafe, W., & Danielewicz, J. (1987). Properties of spoken and written language. In R. Horowitz & S. J. Samuels (Eds.), *Comprehending oral and written language*, 2, 83–113. San Diego: Academic Press.

Cheung, H., & Kemper, S. (1992). Competing complexity metrics and adults' production of complex sentences. *Applied Psycholinguistics*, *13*(01), 53-76.

Corrada, M. M., Brookmeyer, R., Paganini-Hill, A., Berlau, D., & Kawas, C. H. (2010). Dementia incidence continues to increase with age in the oldest old: the 90+ study. *Annals of Neurology*, *67*(1), 114-121.

Craik, F. I., Bialystok, E., & Freedman, M. (2010). Delaying the onset of Alzheimer disease: bilingualism as a form of cognitive reserve. *Neurology*, *75*(19), 1726-1729.

Crossley, S. A., Weston, J. L., McLain Sullivan, S. T., & McNamara, D. S. (2011). The development of writing proficiency as a function of grade level: a linguistic analysis. *Written Communication*, *28*(3), 282-311.

Danner, D. D., Snowdon, D. A., & Friesen, W. V. (2001). Positive emotions in early life and longevity: findings from the Nun Study. *Journal of Personality and Social Psychology*, *80*(5), 804-813.

Farrar, D. E., & Glauber, R. R. (1967). Multicollinearity in regression analysis: the problem revisited. *The Review of Economic and Statistics*, 92-107.

Fillenbaum, G. G., van Belle, G., Morris, J. C., Mohs, R. C., Mirra, S. S., Davis, P. C., Tariot, P. N., Silverman, J. M., Clark, C. M., Welsh-Bohmer, K. A. & Heyman, A. (2008). Consortium to Establish a Registry for Alzheimer's Disease (CERAD): the first twenty years. *Alzheimer's & Dementia*, *4*(2), 96-109.

Fox, N. C., Freeborough, P. A., & Rossor, M. N. (1996). Visualization and quantification of rates of atrophy in Alzheimer's disease. *Lancet*, 348(9020), 94-97.

Fillit, H., Nash, D. T., Rundek, T., & Zuckerman, A. (2008). Cardiovascular risk factors and dementia. *The American Journal of Geriatric Pharmacotherapy*, *6*(2), 100-118.

Grinberg, L. T., & Heinsen, H. (2010). Toward a pathological definition of vascular dementia. *Journal of the Neurological Sciences*, 299(1-2), 136-138.

Hall, C. B., Derby, C., LeValley, A., Katz, M. J., Verghese, J., & Lipton, R. B. (2007). Education delays accelerated decline on a memory test in persons who develop dementia. *Neurology*, *69*(17), 1657-1664.

Hardy, J. A., & Higgins, G. A. (1992). Alzheimer's disease: the amyloid cascade hypothesis. *Science*, *256*(5054), 184.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, *26*(1), 55-88.

Holtzman, D. M., Herz, J., & Bu, G. (2012). Apolipoprotein E and apolipoprotein E receptors: normal biology and roles in Alzheimer disease. *Cold Spring Harbor Perspectives in Medicine*, *2*(3), a006312.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, *45*(3), 337-374.

Hyman, B. T., Phelps, C. H., Beach, T. G., Bigio, E. H., Cairns, N. J., Carrillo, M. C., Dicson, D.W., Duyckaerts, C., Frosch, M. P., Masliah, E., Mirra, S. S., Nelson, P. T., Schneider, J. A., Thal, D. R., Thies, B., Trojanowski, J. Q., Vinters, H. V. & Montine, T. J. (2012). National Institute on Aging–Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimer's & Dementia*, *8*(1), 1-13.

Jefferson, A. L., Gibbons, L. E., Rentz, D. M., Carvalho, J. O., Manly, J., Bennett, D. A., & Jones, R. N. (2011). A life course model of cognitive activities, socioeconomic status, education, reading ability, and cognition. *Journal of the American Geriatrics Society*, *59*(8), 1403-1411.

Kemper, S., Thompson, M., & Marquis, J. (2001). Longitudinal change in language production: effects of aging and dementia on grammatical complexity and propositional content. *Psychology and Aging, 16*, 600-614.

Kemper, S., Greiner, L. H., Marquis, J. G., Prenovost, K., & Mitzner, T. L. (2001) (b). Language decline across the life span: findings from the Nun Study. *Psychology and Aging*, *16*(2), 227-239.

Kintsch, W., & Keenan, J. (1973). Reading rate and retention as a function of the number of propositions in the base structure of sentences. *Cognitive Psychology*, *5*(3), 257-274.

Kleinbaum, D., Kupper, L., Nizam, A., & Rosenberg, E. (2013). Applied regression analysis and other multivariable methods. Boston: Nelson Education.

Kivipelto, M., & Solomon, A. (2008). Alzheimer's disease—the ways of prevention. *The Journal of Nutrition, Health & Aging*, *12*, S89-S94.

Kolb, B., & Whishaw, I. Q. (1998). Brain plasticity and behavior. *Annual Review of Psychology*, *49*(1), 43-64.

Lautenschlager, N. T., Cox, K. L., Flicker, L., Foster, J. K., van Bockxmeer, F. M., Xiao, J., Greenop, K. R. & Almeida, O. P. (2008). Effect of physical activity on cognitive function in older adults at risk for Alzheimer disease: a randomized trial. *Journal of the American Medical Association*, *300*(9), 1027-1037.

Luchsinger, J. A., Tang, M. X., Shea, S., & Mayeux, R. (2004). Hyperinsulinemia and risk of Alzheimer disease. Neurology, 63(7), 1187-1192.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology*, *34*(7), 939-939.

Meng, X., & D'Arcy, C. (2012). Education and dementia in the context of the cognitive reserve hypothesis: a systematic review with meta-analyses and qualitative analyses. *PLoS One*, *7*(6), e38268.

Meng, X., Yu, J., Wang, H., Tan, M., Wang, C., Tan, C., & Tan, L. (2014). Midlife vascular risk factors and the risk of Alzheimer's disease: a systematic review and meta-analysis. *Journal of Alzheimer's Disease*, 42(4), 1295-1310

Meyer, M. R., Tschanz, J. T., Norton, M. C., Welsh-Bohmer, K. A., Steffens, D. C., Wyse, B. W., & Breitner, J. C. (1998). *APOE* genotype predicts when—not whether—one is predisposed to develop Alzheimer disease. *Nature Genetics*, *19*(4), 321-322.

Mirra, S. S., Heyman, A., McKeel, D., Sumi, S. M., Crain, B. J., Brownlee, L. M., Vogel, F. S., Hughes, J. P., van Belle, G., Berg, L. (1991). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part II. standardization of the neuropathologic assessment of Alzheimer's disease. *Neurology*, *41*(4), 479-486.

Mitzner, T. L., & Kemper, S. (2003). Oral and written language in late adulthood: findings from the Nun Study. *Experimental Aging Research*, *29*(4), 457-474.

Morris, M. C., Tangney, C. C., Wang, Y., Sacks, F. M., Bennett, D. A., & Aggarwal, N. T. (2015). MIND diet associated with reduced incidence of Alzheimer's disease. *Alzheimer's & Dementia*, *11*(9), 1007-1014.

Mueller, K. D., Koscik, R. L., LaRue, A., Clark, L. R., Hermann, B., Johnson, S. C., & Sager, M. A. (2015). Verbal fluency and early memory decline: results from the Wisconsin registry for Alzheimer's prevention. *Archives of Clinical Neuropsychology*, *30*(5), 448-457.

Mummery, C. J., Patterson, K., Price, C. J., Ashburner, J., Frackowiak, R. S. J., & Hodges, J. R. (2000). A voxel-based morphometry study of semantic dementia: relationship between temporal lobe atrophy and semantic memory. *Annals of Neurology*, *47*(1), 36-45.

National Institute on Aging. (2016). Preventing Alzheimer’s disease: what do we know?. Retrieved from: https://www.nia.nih.gov/alzheimers/publication/preventing-alzheimers-disease/risk-factors-alzheimers-disease

Norton, S., Matthews, F. E., Barnes, D. E., Yaffe, K., & Brayne, C. (2014). Potential for primary prevention of Alzheimer's disease: an analysis of population-based data. *The Lancet Neurology*, *13*(8), 788-794.

Pakhomov, S. V., & Hemmy, L. S. (2014). A computational linguistic measure of clustering behavior on semantic verbal fluency task predicts risk of future dementia in the Nun Study. *Cortex*, *55*, 97-106.

Patzwald, G. A., & Wildt, S. (2004). The use of convent archival records in medical research: the School Sisters of Notre Dame archives and the Nun Study. *The American Archivist*, *67*(1), 86-106.

Perquin, M., Vaillant, M., Schuller, A. M., Pastore, J., Dartigues, J. F., Lair, M. L., Diederich, N. & MemoVie Group. (2013). Lifelong exposure to multilingualism: new evidence to support cognitive reserve hypothesis. *PloS One*, *8*(4), e62030.

Prince, M., Bryce, R., Albanese, E., Wimo, A., Ribeiro, W., & Ferri, C. P. (2013). The global prevalence of dementia: a systematic review and meta-analysis. *Alzheimer's & Dementia*, *9*(1), 63-75.

Prince, M., Albanese, E., Guerchet, M. & Prina, M. (2014). World Alzheimer report: dementia and risk reduction: an analysis of protective and modifiable factors. London. *Alzheimer's Disease International.*

Orgogozo, J. M., Dartigues, J. F., Lafont, S., Letenneur, L., Commenges, D., Salamon, R., Renaud, S. & Breteler, M. B. (1997). Wine consumption and dementia in the elderly: a prospective community study in the Bordeaux area. *Revue Neurologique*, *153*(3), 185-192.

Ravid, D., & Berman, R. A. (2006). Information density in the development of spoken and written narratives in English and Hebrew. *Discourse Processes*, *41*(2), 117-149.

Riley, K. P., Snowdon, D. A., Desrosiers, M. F., & Markesbery, W. R. (2005). Early life linguistic ability, late life cognitive function, and neuropathology: findings from the Nun Study. *Neurobiology of Aging*, *26* (3), 341-347.

Rusanen, M., Kivipelto, M., Quesenberry, C. P., Zhou, J., & Whitmer, R. A. (2011). Heavy smoking in midlife and long-term risk of Alzheimer disease and vascular dementia. *Archives of Internal Medicine*, *171*(4), 333-339.

SantaCruz, K. S., Sonnen, J. A., Pezhouh, M. K., Desrosiers, M. F., Nelson, P. T., & Tyas, S. L. (2011). Alzheimer disease pathology in subjects without dementia in 2 studies of aging: the Nun Study and the Adult Changes in Thought Study. *Journal of Neuropathology & Experimental Neurology*, *70*(10), 832-840.

SAS Institute Inc. (2017). The LOGISTIC procedure – regression diagnostics. Retrieved from: http://support.sas.com/documentation/cdl/en/statug/68162/HTML/default/viewer.htm#statug_logistic_details38.htm

SAS Institute Inc. (2017b). SAS/STAT® 9.2 User's Guide, Second Edition. Firth's penalized likelihood compared with other approaches. Retrieved from: https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_logistic_sect063.htm

Saykin, A. J., Wishart, H. A., Rabin, L. A., Santulli, R. B., Flashman, L. A., West, J. D., McHugh, MA & Mamourian, A. C. (2006). Older adults with cognitive complaints show brain atrophy similar to that of amnestic MCI. *Neurology*, *67*(5), 834-842.

Selkoe, D. J. (2001). Alzheimer's disease: genes, proteins, and therapy. *Physiological reviews*, *81*(2), 741-766.

Snowdon, D. A., Kemper, S. J., Mortimer, J. A., Greiner, L. H., Wekstein, D. R., & Markesbery, W. R. (1996). Linguistic ability in early life and cognitive function and Alzheimer's disease in late life: findings from the Nun Study. *Journal of the American Medical Association*, *275*(7), 528-532.

Snowdon, D. A., Greiner, L. H., Mortimer, J. A., Riley, K. P., Greiner, P. A., & Markesbery, W. R. (1997). Brain infarction and the clinical expression of Alzheimer disease: the Nun Study. *Journal of the American Medical Association, 277*(10), 813-817.

Snowdon, D. A. (1997). Aging and Alzheimer's disease: lessons from the Nun Study. *The Gerontologist*, *37*(2), 150-156.

Snowdon, D. A., Greiner, L. H., & Markesbery, W. R. (2000). Linguistic ability in early life and the neuropathology of Alzheimer's disease and cerebrovascular disease: Findings from the Nun Study. *Annals of the New York Academy of Sciences*, *903*(1), 34-38

Stern, Y., Gurland, B., Tatemichi, T. K., Tang, M. X., Wilder, D., & Mayeux, R. (1994). Influence of education and occupation on the incidence of Alzheimer's disease. *Journal of the American Medical Association*, *271*(13), 1004-1010.

Stern, Y. (2002). What is cognitive reserve? theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*, *8*(03), 448-460.

Stern, Y. (2012). Cognitive reserve in ageing and Alzheimer's disease. *The Lancet Neurology*, *11*(11), 1006-1012.

Szklo, M. (1998). Population-based cohort studies. *Epidemiologic Reviews*, *20*(1), 81-90.

Tanzi, R. E., & Bertram, L. (2005). Twenty years of the Alzheimer's disease amyloid hypothesis: A genetic perspective. *Cell*, *120*(4), 545-555.

Tyas, S. L., Koval, J. J., & Pederson, L. L. (2000). Does an interaction between smoking and drinking influence the risk of Alzheimer's disease? results from three Canadian data sets. *Statistics in Medicine*, *19*(11-12), 1685-1696.

Tyas, S. L. (2001). Alcohol use and the risk of developing Alzheimer's disease. *Alcohol Research and Health*, *25*(4), 299-307.

Tyas, S. L., White, L. R., Petrovitch, H., Ross, G. W., Foley, D. J., Heimovitz, H. K., & Launer, L. J. (2003). Mid-life smoking and late-life dementia: the Honolulu-Asia Aging Study. *Neurobiology of Aging*, *24*(4), 589-596.

Tyas, S. L., & Gutmanis, I. (2015). Alzheimer's disease. *Managerial epidemiology: concepts and cases, 3rd Edition*., 467-508. Chicago, Illinois. Health Administration Press.

Waldemar, G., Dubois, B., Emre, M., Georges, J., McKeith, I. G., Rossor, M., Scheltens, P., Tariska, P. & Winblad, B. (2007). Recommendations for the diagnosis and management of Alzheimer's disease and other disorders associated with dementia: EFNS guideline. *European Journal of Neurology*, *14*(1), e1-e26.

World Health Organization. (2014). The world health report 2008: primary health care (now more than ever). Retrieved from: http://www.who.int/whr/2008/en/

World Health Organization (2015). Ageing and health. Retrieved from: http://www.who.int/mediacentre/factsheets/fs404/en/

World Health Organization. (2017). Dementia. Retrieved from: http://www.who.int/mediacentre/factsheets/fs362/en/

Wu, L., Rosa-Neto, P., Hsiung, G. Y. R., Sadovnick, A. D., Masellis, M., Black, S. E., Jia, J. & Gauthier, S. (2012). Early-onset familial Alzheimer's disease (EOFAD). *The Canadian Journal of Neurological Sciences*, *39*(04), 436-445.

Yeung, C. M., John, P. D. S., Menec, V., & Tyas, S. L. (2014). Is bilingualism associated with a lower risk of dementia in community-living older adults? cross-sectional and prospective analyses. *Alzheimer Disease & Associated Disorders*, *28*(4), 326-332.

Appendix

Appendix A: Literature Search Strategies

Table A1. Literature Search Strategy: PubMed

	Search Strategy #1			
Database:	**Early Life including Childhood**	**Written Language Skills**	**Late Life**	**Cognitive Outcomes**
PubMed/Medline	Early life[tiab] OR Youth[tiab] OR Young[tiab] OR Adolescent[MeSH] OR Infant[MeSH] OR Infan*[tiab] OR Child[MeSH] OR Child*[tiab]	Writing[MeSH] OR Writing[tiab] OR Handwriting[MeSH] OR Handwriting[tiab] OR language*[tiab] OR multilingual*[tiab] OR bilingual*[tiab] OR language[MeSH]	Aged [MeSH] OR Aging [MeSH] Older adult [tiab] OR Later life [tiab] OR Life outcomes [tiab] OR Elderly [tiab] OR Seniors [tiab]	Cognitive reserve [MeSH] OR Cognitive reserve [all fields] OR Brain reserve [all fields] OR Alzheimer Disease [MeSH] OR Alzheimer disease [tiab] OR Dementia [MeSH] OR Dementia [tiab] OR Cognitive resilience [all fields] OR Cognitive decline [all fields]

Overall search strategy: #1 AND #2 AND #3 AND #4
(Retrieved 341 records)

#4 Cognitive reserve[MeSH] OR cognitive reserve[all fields] OR brain reserve[all fields] OR Alzheimer disease[MeSH] OR Alzheimer disease[tiab] OR Dementia[MeSH] OR Dementia[tiab] OR cognitive resilience[all fields] OR cognitive decline[all fields]
#3 Aged[MeSH] OR Aging[MeSH] OR older adult[tiab] OR later life[tiab] OR life outcomes[tiab] OR elderly[tiab] OR seniors[tiab]
#2 Writing[MeSH] OR writing[tiab] OR handwriting[MeSH] OR handwriting[tiab] OR language*[tiab] or multilingual*[tiab] OR bilingual*[tiab] OR language[MeSH]
#1 Early life[tiab] OR youth[tiab] OR young[tiab] OR adolescent[MeSH] OR Infant[MeSH] OR Infan*[tiab] OR Child[MeSH] OR Child*[tiab]

Table A2. Literature Search Strategy: PsycINFO

	Search Strategy #2			
Concept:	**Early Life including Childhood**	**Written Language Skills**	**Late Life**	**Cognitive Outcomes**
Author Keyword	Early life Youth Child	Written language Writing skills Writing Handwriting	Aging Geriatrics Aged Later life Elderly	Alzheimer* Dementia Cognitive decline Cognitive function Cognitive res* Cognitive reserve
Subject Headings Index Terms	Early experience Childhood development Infant development	Written language Writing skills Linguistics Language Oral communication Written communication	Aging	Alzheimer's disease Dementia Cognitive impairment Cognitive decline Cognitive ability

Overall search strategy: #1 AND #2 AND #3 AND #4
(Retrieved 201 records)

#4 Alzheimer* OR Dementia OR "Cognitive decline" OR "Cognitive function" OR "Cognitive res*" OR "Cognitive reserve" OR Alzheimer's disease OR "Cognitive impairment" OR "Cognitive ability"
#3 Aging OR Geriatrics OR Aged OR "Later life" OR Elderly
#2 "Written language" OR "Writing skills" OR Writing OR Handwriting OR Linguistics OR Language OR "Oral communication" OR "Written communication"
#1 "Early life" OR Youth OR "Early experience" OR "Childhood development" or "Infant development"

Appendix B: Literature Summary Table

Table B1: Summary Table for Findings on the Association Between Early-life Written Language Skills and Later-life Cognitive Outcomes

Study	Study Population, Sample Characteristics	Exposure and Covariates	Outcome	Analysis	Results
Brewster et al., 2014 Life experience and demographic influences on cognitive function in older adults	Participants included 333 diverse older adults from the UC Davis Aging Diversity Cohort (a longitudinal study of cognitive aging).	Exposure measures included baseline episodic memory, executive function, semantic memory and life experience (i.e., literacy, early socioeconomic status, life course physical and recreational activity measures). Covariates included *APOE*, age, and demographic characteristics (such as education, ethnicity, and language).	Cognitive function in later life was the outcome of interest. Cognitive outcomes were determined using a composite measure of memory (episodic and semantic), from the Spanish and English Neuropsychological Assessment Scales.	To analyze the data, ANOVA, chi-square tests, and mixed effects regression analyses were utilized.	Early factors such as beneficial SES were associated with less cognitive decline. Bilingual participants did not significantly differ from monolingual speakers in terms of longitudinal cognitive outcomes. At baseline, reading had strong effects on cognition in older adults.
Cheung and Kemper, 1992 Competing	Language samples were obtained from 30 different adults,	The exposure was verbal ability and aspects of language usage with	The outcome was understanding the usefulness of complexity metrics.	Analyses such as ANOVA, goodness of fit measures, covariance matrices	Most of the language metrics provided adequate descriptions of

Study	Study Population, Sample Characteristics	Exposure and Covariates	Outcome	Analysis	Results
complexity metrics and adults' production of complex sentences.	between the ages of 60-90.	covariates including educational level, age and vocabulary.	The outcome was measured using complexity metrics that measured aspects of language such as content, length, amount of embedding, type of embedding, and complexity in language samples.	and chi-square were used.	individual differences and age-related changes in complexity. Embedding was helpful in predicting how easily sentences could be understood or recalled.
Craik et al., 2010 Delaying the onset of Alzheimer disease: bilingualism as a form of cognitive reserve	Data from 211 individuals diagnosed with probable AD from the Sam and Ida Ross Memory Clinic were utilized.	The exposure was bilingualism as measured by having spent the majority of their lives, at least from early adulthood using a minimum of two languages. Covariates included education, occupational history, place of	The outcome of interest was cognitive reserve based on diagnosis of probable AD.	Measures of odds ratios, 95% confidence intervals, means, and logistic regression models with adjustments for confounding variables were used for analysis.	Bilingual patients were diagnosed 4.3 years later and symptoms tended to appear later (5.1 years) in comparison to monolingual individuals.

Study	Study Population, Sample Characteristics	Exposure and Covariates	Outcome	Analysis	Results
		birth, and immigration status.			
Jefferson et al., 2011 A life course model of cognitive activities, socioeconomic status, education, reading ability, and cognition	The sample utilized 951 older adults who were free of dementia at baseline from the Rush and Aging project. A subset of this larger population (n=260) met the criteria for possible dementia and were used for this analysis.	Exposure measures included life course factors such as early-, mid- and late-life participation in cognitive activities, socioeconomic status in early life and adulthood, education and reading ability. Covariates included age, gender, and race.	The outcome was cognitive reserve and beneficial states of late-life cognition (prior to the onset of dementia).	A composite measure was created by combining indicators and scores of early-, mid- and late-life activity. Path analysis, descriptive analyses (e.g., means, percentages, and standard deviations), and priori modelling techniques were used for analysis.	Education showed the strongest association with late-life cognition. Reading ability was strongly associated with working memory, episodic memory and global cognition.
Mitzner & Kemper, 2003 Oral and written language in late adulthood: findings	Written and oral samples from 118 women were used for analysis.	The exposure was language samples either oral or written. Oral samples were taken from short autobiographies	The outcome was participant characteristics and cognitive function based on MMSE and ADL scores as representations of	For analysis, ranges, means, standard deviations, t-tests, and variance and correlations were determined.	Participant characteristics throughout life (e.g., education, cognitive status, and physical function)

Study	Study Population, Sample Characteristics	Exposure and Covariates	Outcome	Analysis	Results
from the Nun Study		provided following assessments. Written samples were taken from written autobiographies.	cognitive status and physical function.		influenced grammatical and conceptual characteristics of language samples. Written language samples have greater power than oral language samples to differentiate between ability levels in older adults.
Mueller et al., 2015 Verbal fluency and early memory decline: results from the Wisconsin registry for Alzheimer's prevention	The sample population involved 283 participants from a longitudinal cohort study (WRAP).	The exposure of interest was measures of verbal fluency variables using tests such as Controlled Oral Word Association Test. Covariates included age, gender, and literacy level.	The outcome of interest was cognitive states (i.e., cognitive healthy, amnestic form of mild cognitive impairment (aMCI).	Analyses included ANOVA, t-tests, significance testing, and chi-square.	Lower scores in verbal fluency tests were obtained in individuals with aMCI in comparison to cognitively healthy individuals.

Study	Study Population, Sample Characteristics	Exposure and Covariates	Outcome	Analysis	Results
Perquin et al., 2013 Lifelong exposure to multilingualism: new evidence to support cognitive reserve hypothesis	The study population involved 232 non-demented participants.	The exposure was multilingual ability measured by the number of fluent languages practiced throughout life. Confounding variables such as gender, and sociocultural factors were addressed.	The outcome was cognitive outcomes (e.g., cognitive impairment no dementia (CIND) or CIND-free).	Confidence intervals, odds ratios, p-values, significance testing and logistic regression models were used for analyses.	Individuals who practiced more than 2 languages throughout life were 3 times less likely to have CIND in later life.
Riley et al., 2005 Early life linguistic ability, late life cognitive function, and neuropathology: findings from the Nun Study	The overall population involved 678 Nun Study participants aged 75 to 102. However, a subset of 90 participants was used for further analyses.	The exposure was a measure of linguistic ability using idea density (derived from handwritten autobiographies).	The outcome was cognitive state based on cognitive function (including CERAD battery of neuropsychological tests and activities of daily living measures) and AD neuropathology. AD neuropathology was determined through the	Logistic regression, significance testing and P-values were determined for analyses using SAS statistical software package.	Low idea density increased as the severity of cognitive impairment. Significant relationships also existed between low idea density and mild cognitive impairments. MMSE and

Study	Study Population, Sample Characteristics	Exposure and Covariates	Outcome	Analysis	Results
			assessment of neuropathologic lesions using Braak staging method.		Delayed Word Recall tests were the best predictors of idea density.
Snowdon et al., 1996 Linguistic ability in early life and cognitive function and Alzheimer's disease in later life	The study population comprised of 678 Nun Study participants. A subset of the Nun Study population including 93 participants who were between the ages of 75 to 95 years at time of assessment were used.	The exposure of interest was linguistic ability as measured by idea density and grammatical complexity from scored autobiographies. Covariates included age at the time of functional assessment, and years of education.	The outcome was diagnosis of Alzheimer's disease based on 7 tests of cognitive function (e.g., Delayed Word Recall, Verbal Fluency, Boston Naming, MMSE, etc.), and neuropathologic evaluations (measures of senile plaques and tangles).	For analysis, means, standard deviations, Pearson correlation coefficients, odds ratios and confidence intervals were utilized.	Significant associations between idea density, grammatical complexity ($P<0.001$) and years of education ($P<0.01$) were determined. Idea density had the strongest association with cognitive function. Those with low idea density had substantially more

Study	Study Population, Sample Characteristics	Exposure and Covariates	Outcome	Analysis	Results
					NFT than those with higher idea density.
Snowdon et al., 2000 Linguistic ability in early life and the neuropathology of Alzheimer's disease and cerebrovascular disease	The study population comprises participants from the Nun Study, ages 75 to 102. The original sample included 678 participants. A subset of 74 participants were used for this analysis.	Indicators of linguistic ability: idea density and grammatical complexity were identified. However, only idea density was used for this analysis. Covariates included age at death, and location of convent.	The outcome was measured based on neuropathologic criteria and included whether or not brain infarcts were identified, NFTs in certain areas of the brain were identified, and whether or not the participant met neuropathologic criteria for AD based.	For analysis, means, idea density scores, 95% confidence intervals, and significance testing were utilized.	Those who met neuropathologic criteria for AD tended to have lower idea density scores in comparison to those who did not meet the criteria. Presence of NFT in the brain was strongly associated with lower idea density scores. No consistent associations between idea density and infarcts were determined.

Appendix C: Description of Grammatical Complexity and Idea Density

Table C1: Levels of Grammatical Complexity

Level	Description of Complexity Level
Level 0	"Simple, one-clause sentences" (Kemper et al., 1992) ***Example of GC level 0 from a Milwaukee sister:*** "Two of the boys are dead." (Snowdon et al., 1996).
Level 1	"Complex sentences with embedded infinitival complements" (Kemper et al., 1992)
Level 2	"Complex sentences with *wh*-predicate[1] complements, conjoined clauses, and compound subjects" (Kemper et al., 1992) [1]*Wh*- predicate refers to words beginning with *wh*- (e.g., what, where, etc.)
Level 3	"Complex sentences with relative clauses modifying the object noun phrase or with predicate noun phrase complements" (Kemper et al., 1992)
Level 4	"Complex sentences with gerundive complements or comparative constructions" (Kemper et al., 1992)
Level 5	"Complex sentences with relative clauses modifying the subject noun phrase, subject noun phrase complements, and subject nominalizations" (Kemper et al., 1992) ***Example of GC level 5 from a Milwaukee sister:*** "I prefer teaching music to any other profession" (Snowdon et al., 1996)
Level 6	"Complex sentences with subordinate clauses" (Kemper et al., 1992)
Level 7	"Complex sentences with multiple forms of embedding and subordination" (Kemper et al., 1992) ***Example of GC level 7 from a Milwaukee sister:*** "The happiest day of my life so far was my First Communion Day which was in June nineteen hundred and twenty when I was but eight years of age, and four years later in the same month I was confirmed by Bishop D. D. McGavick." (Snowdon et al., 1996)

Table C2: Computing Idea Density

Example Sentence from an Autobiography	**Computation of Idea Density**
"I was born in Eau Claire, Wis, on May 24, 1913 and was baptized in St James Church." (Snowdon et al., 1996) *(Idea density score: 3.9 per 10 words)*	**List of Ideas:** (Snowdon et al., 1996) 1. "I was born" 2. "born in Eau Claire Wis" 3. "born on May 24, 1913" 4. "I was baptized" 5. "was baptized in church" 6. "was baptized in St James Church 7. "I was born…and was baptized." **Calculation:** • There were a total of 7 ideas (listed above). The 7 ideas are then divided by the total number of words in the sentence (18 words) and multiplied by 10 to provide an idea density score of 3.9 per 10 words (Snowdon et al., 1996).

Note: *Idea density was calculated by determining the average number of ideas expressed per 10 words (Kemper et al., 2001b).*

Appendix D: Sensitivity Analyses Using a Sample of Participants that Met Either CERAD or NIA-RI Criteria

In the present study, results were impacted by small sample sizes. The NIA-RI sample was particularly impacted. When running logistic regression models using the NIA-RI sample, the relationship between idea density and cognitive resilience could not be completed. In the NIA-RI sample, no participants with low idea density reflected cognitive resilience. In addition, a significant relationship was not found between grammatical complexity and cognitive resilience in the NIA-RI sample. Small sample sizes and challenges with the NIA-RI sample specifically affected questions two, three and four.

In an effort to address these issues, the coding of CERAD and NIA-RI was modified in sensitivity analyses. As indicated in the analytic sample flow chart (Figure 1), the NIA-RI analytic sample is a subsample of the CERAD analytic sample. In the sensitivity analyses, instead of separating the CERAD and NIA-RI samples, a combined variable using coding that allowed for either CERAD or NIA-RI criteria to be met was derived. The analytic sample with new coding increased the separate sample sizes from CERAD (n=56) and NIA-RI (n=42) to a larger overall sample including either CERAD or NIA-RI (n=62). Tables summarizing results from these sensitivity analyses are provided in Appendix D.

Sensitivity analyses considered research questions two, three and four: whether the association between written language skills and cognitive resilience held when adjusting for standard covariates (age and *APOE*), whether the relationship held when adjusting for an indicator of cognitive reserve (education), and indicators of brain reserve (cerebral infarcts and cortical atrophy). Research question one focused on descriptive analyses. Sensitivity analyses were not relevant for question one since there were no issues associated with sample size.

Sensitivity analyses using the combined CERAD or NIA-RI criteria (n=62) provided more consistent findings in comparison to the CERAD sample (n=56). Previous literature predominantly analyzed the two samples (CERAD or NIA-RI) separately. However, using a larger combined sample, idea density and grammatical complexity were consistent significant predictors of cognitive resilience in both unadjusted and full models. These findings suggested that inconsistencies for idea density in the CERAD sample may have been influenced by the lower power of the small sample size.

Table D1: The Association Between Idea Density and Cognitive Resilience, Either CERAD or NIA-RI Criteria (n=62)

Cognitive Resilience, Either CERAD or NIA-RI Criteria (n=62)

	Unadjusted Model	Age	*APOE*-ε4 Presence	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable				
Idea Density	**0.09**	**0.08**	**0.09**	**0.09**
Low[1] vs. high	**(0.01-0.35)**	**(0.01-0.35)**	**(0.01-0.39)**	**(0.01-0.39)**
Age (years)		0.98 (0.82-1.17)		0.98 (0.82-1.18)
APOE-ε4 Presence			0.66 (0.19-2.24)	0.66 (0.19-2.24)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **NIA-RI** = National Institute of Aging – Reagan Institute; **OR** = Odds ratio

Table D2: The Association Between Grammatical Complexity and Cognitive Resilience, Either CERAD or NIA-RI Criteria (n=62)

Cognitive Resilience, Either CERAD or NIA-RI Criteria (n=62)

	Unadjusted Model	Age	*APOE*-ε4 Presence	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable				
Grammatical Complexity *Low*[1] *vs. high*	**0.14 (0.03-0.48)**	**0.14 (0.03-0.49)**	**0.12 (0.03-0.44)**	**0.12 (0.03-0.45)**
Age (years)		1.02 (0.86-1.21)		1.01 (0.85-1.21)
APOE-ε4 Presence			0.39 (0.11-1.26)	0.39 (0.11-1.26)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **NIA-RI** = National Institute of Aging – Reagan Institute; **OR** = Odds ratio

Table D3: The Impact of Education on the Association Between Idea Density and Cognitive Resilience Using Firth Logistic Regression, Either CERAD or NIA-RI Criteria (n=62)

Cognitive Resilience, Either CERAD or NIA-RI Criteria (n=62)

	Unadjusted Model	Age	*APOE*-ε4 Presence	Age and *APOE*	Education	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable						
Idea density	**0.11**	**0.11**	**0.12**	**0.12**	**0.12**	**0.12**
Low[1] *vs. high*	**(0.02-0.47)**	**(0.02-0.49)**	**(0.03-0.53)**	**(0.03-0.55)**	**(0.03-0.53)**	**(0.03-0.59)**
Age (years)		0.98 (0.83-1.16)		0.98 (0.83-1.17)		0.94 (0.78-1.13)
APOE-ε4 Presence			0.67 (0.20-2.22)	0.67 (0.02-2.22)		0.57 (0.17-1.95)
Education ≤ *High school vs. Bachelor's degree*					0.09 (0.003-3.14)	0.06 (0.002-2.44)
Master's vs. Bachelor's degree					0.65 (0.21-2.06)	0.59 (0.18-2.00)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **NIA-RI** = National Institute of Aging – Reagan Institute; **OR** = Odds ratio

Table D4: The Impact of Education on the Association Between Grammatical Complexity and Cognitive Resilience Using Firth Logistic Regression, Either CERAD or NIA-RI Criteria (n=62)

Cognitive Resilience, Either CERAD or NIA-RI Criteria (n=62)

	Unadjusted Model	Age	*APOE*-ε4 Presence	Age and *APOE*	Education	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable						
Grammatical Complexity *Low*[1] *vs. high*	**0.15 (0.04-0.58)**	**0.16 (0.04-0.62)**	**0.14 (0.04-0.55)**	**0.15 (0.04-0.58)**	**0.21 (0.05-0.78)**	**0.19 (0.05-0.76)**
Age (years)		1.01 (0.86-1.20)		1.01 (0.85-1.20)		0.98 (0.82-1.17)
APOE-ε4 Presence			0.41 (0.13-1.34)	0.42 (0.13-1.36)		0.38 (0.12-1.25)
Education ≤ *High school vs. Bachelor's degree*					0.15 (0.01-4.79)	0.11 (0.003-3.94)
Master's vs. Bachelor's degree					0.76 (0.25-2.29)	0.73 (0.22-2.37)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **NIA-RI** = National Institute of Aging – Reagan Institute; **OR** = Odds ratio

Table D5: The Association Between Idea Density and Cognitive Resilience Adjusted for Presence of Cerebral Infarcts, Either CERAD or NIA-RI Criteria (n=61)

Cognitive Resilience, Either CERAD or NIA-RI Criteria (n=61)

	Unadjusted Model	Age	*APOE*-ε4 Presence	Presence of Infarcts	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable					
Idea Density *Low*[1] *vs. high*	**0.09 (0.01-0.37)**	**0.08 (0.01-0.35)**	**0.10 (0.01-0.40)**	**0.10 (0.01-0.42)**	**0.10 (0.01-0.47)**
Age (years)		0.95 (0.79-1.15)			0.98 (0.80-1.19)
APOE-ε4 Presence			0.68 (0.20-2.35)		0.68 (0.19-2.34)
Presence of Infarcts (yes vs. no)				0.57 (0.16-1.99)	0.58 (0.15-2.18)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **NIA-RI** = National Institute of Aging – Reagan Institute; **OR** = Odds ratio

Table D6: The Association Between Grammatical Complexity and Cognitive Resilience Adjusted for Presence of Cerebral Infarcts, Either CERAD or NIA-RI Criteria (n=61)

Cognitive Resilience, Either CERAD or NIA-RI Criteria (n=61)

	Unadjusted Model	Age	*APOE*-ε4 Presence	Presence of Infarcts	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable					
Grammatical Complexity *Low*[1] *vs. high*	**0.14 (0.03-0.50)**	**0.14 (0.03-0.50)**	**0.13 (0.03-0.47)**	**0.15 (0.03-0.54)**	**0.13 (0.03-0.51)**
Age (years)		0.99 (0.83-1.19)			1.03 (0.85-1.26)
APOE-ε4 Presence			0.41 (0.12-1.32)		0.39 (0.11-1.29)
Presence of Infarcts *(yes vs. no)*				0.45 (0.13-1.5)	0.41 (0.10-1.47)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **NIA-RI** = National Institute of Aging – Reagan Institute; **OR** = Odds ratio

Table D7: The Association Between Idea Density and Cognitive Resilience Adjusted for Number of Cerebral Infarcts, Either CERAD or NIA-RI Criteria (n=61)

Cognitive Resilience, Either CERAD or NIA-RI Criteria (n=61)

	Unadjusted Model	Age	*APOE*-ε4 Presence	Number of Infarcts	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable					
Idea Density *Low*[1] *vs. high*	**0.09 (0.01-0.37)**	**0.08 (0.01-0.35)**	**0.10 (0.01-0.40)**	**0.09 (0.01-0.40)**	**0.09 (0.01-0.42)**
Age (years)		0.95 (0.79-1.15)			0.96 (0.79-1.17)
APOE-ε4 Presence			0.68 (0.20-2.35)		0.69 (0.20-2.37)
Number of Infarcts				0.95 (0.61-1.40)	0.97 (0.62-1.45)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **NIA-RI** = National Institute of Aging – Reagan Institute; **OR** = Odds ratio

Table D8: The Association Between Grammatical Complexity and Cognitive Resilience Adjusted for Number of Cerebral Infarcts, Either CERAD or NIA-RI Criteria (n=61)

Cognitive Resilience, Either CERAD or NIA-RI Criteria (n=61)

	Unadjusted Model	Age	*APOE*-ε4 Presence	Number of Infarcts	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable					
Grammatical Complexity *Low*[1] *vs. high*	**0.14 (0.03-0.50)**	**0.14 (0.03-0.50)**	**0.13 (0.03-0.47)**	**0.15 (0.03-0.53)**	**0.13 (0.03-0.49)**
Age (years)		0.99 (0.83-1.19)			1.01 (0.84-1.22)
APOE-ε4 Presence			0.41 (0.18-1.32)		0.41 (0.12-1.32)
Number of Infarcts				0.87 (0.57-1.25)	0.87 (0.56-1.27)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **NIA-RI** = National Institute of Aging – Reagan Institute; **OR** = Odds ratio

Table D9: The Association Between Idea Density and Cognitive Resilience Adjusted for Cortical Atrophy, Either CERAD or NIA-RI Criteria (n=58)

Cognitive Resilience, Either CERAD or NIA-RI Criteria (n=58)

	Unadjusted Model	Age	*APOE*-ε4 Presence	Cortical Atrophy	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable					
Idea Density	**0.09**	**0.09**	**0.10**	**0.10**	**0.10**
Low[1] vs. high	**(0.01-0.38)**	**(0.01-0.37)**	**(0.02-0.43)**	**(0.01-0.43)**	**(0.01-0.46)**
Age (years)		0.97 (0.80-1.17)			0.96 (0.79-1.16)
APOE-ε4 Presence			0.58 (0.16-2.05)		0.59 (0.16-2.17)
Cortical Atrophy *(yes vs. no)*				0.41 (0.11-1.37)	0.40 (0.11-1.36)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **NIA-RI** = National Institute of Aging – Reagan Institute; **OR** = Odds ratio

Table D10: The Association Between Grammatical Complexity and Cognitive Resilience Adjusted for Cortical Atrophy, Either CERAD or NIA-RI Criteria (n=58)

Cognitive Resilience, Either CERAD or NIA-RI Criteria (n=58)

	Unadjusted Model	Age	*APOE*-ε4 Presence	Cortical Atrophy	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable					
Grammatical Complexity *Low[1] vs. high*	**0.16 (0.03-0.59)**	**0.16 (0.03-0.60)**	**0.15 (0.03-0.56)**	**0.18 (0.04-0.67)**	**0.16 (0.03-0.63)**
Age (years)		1.00 (0.84-1.20)			0.99 (0.82-1.20)
APOE-ε4 Presence			0.37 (0.10-1.22)		0.38 (0.10-1.28)
Cortical Atrophy *(yes vs. no)*				0.41 (0.12-0.31)	0.41 (0.12-1.38)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **NIA-RI** = National Institute of Aging – Reagan Institute; **OR** = Odds ratio

Table D11: The Association Between Idea Density and Cognitive Resilience Adjusted for Presence of Cerebral Infarcts and Cortical Atrophy, Either CERAD or NIA-RI Criteria (n=58)

Cognitive Resilience, Either CERAD or NIA-RI Criteria (n=58)

	Unadjusted Model	Age	*APOE*-ε4 Presence	Presence of Infarcts	Cortical Atrophy	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable						
Idea Density *Low[1] vs. high*	**0.09 (0.01-0.38)**	**0.09 (0.01-0.37)**	**0.10 (0.02-0.43)**	**0.11 (0.02-0.45)**	**0.10 (0.01-0.43)**	**0.13 (0.02-0.60)**
Age (years)		0.97 (0.80-1.17)				0.99 (0.81-1.21)
APOE-ε4 Presence			0.58 (0.16-2.05)			0.59 (0.15-21.18)
Presence of Infarcts *(yes vs. no)*				0.46 (0.12-1.70)		0.40 (0.09-1.64)
Cortical Atrophy *(yes vs. no)*					0.41 (0.11-1.37)	0.36 (0.09-1.28)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **NIA-RI** = National Institute of Aging – Reagan Institute; **OR** = Odds ratio

Table D12: The Association Between Grammatical Complexity and Cognitive Resilience Adjusted for Presence of Cerebral Infarcts and Cortical Atrophy, Either CERAD or NIA-RI Criteria (n=58)

Cognitive Resilience, Either CERAD or NIA-RI Criteria (n=58)

	Unadjusted Model	Age	*APOE*-ε4 Presence	Presence of Infarcts	Cortical Atrophy	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable						
Grammatical complexity *Low[1] vs. high*	**0.16 (0.03-0.59)**	**0.16 (0.03-0.60)**	**0.15 (0.03-0.56)**	**0.18 (0.04-0.66)**	**0.18 (0.04-0.67)**	**0.18 (0.03-0.73)**
Age (years)		1.00 (0.84-1.20)				1.03 (0.84-1.27)
APOE-ε4 Presence			0.37 (0.10-1.22)			0.38 (0.10-1.34)
Presence of Infarcts *(yes vs. no)*				0.39 (0.10-1.34)		0.32 (0.07-1.23)
Cortical Atrophy *(yes vs. no)*					0.41 (0.12-1.31)	0.38 (0.10-1.31)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **NIA-RI** = National Institute of Aging – Reagan Institute; **OR** = Odds ratio

Table D13: The Association Between Idea Density and Cognitive Resilience Adjusted for Number of Cerebral Infarcts and Cortical Atrophy, Either CERAD or NIA-RI Criteria (n=58)

Cognitive Resilience, Either CERAD or NIA-RI Criteria (n=58)

	Unadjusted Model	Age	*APOE*-ε4 Presence	Number of Infarcts	Cortical Atrophy	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable						
Idea Density	**0.09**	**0.09**	**0.10**	**0.10**	**0.10**	**0.10**
Low[1] vs. high	**(0.01-0.38)**	**(0.01-0.37)**	**(0.02-0.43)**	**(0.01-0.42)**	**(0.01-0.43)**	**(0.01-0.50)**
Age (years)		0.97 (0.80-1.17)				0.96 (0.79-1.17)
APOE-ε4 Presence			0.58 (0.16-2.05)			0.59 (0.16-2.17)
Number of Infarcts				0.92 (0.58-1.37)		0.96 (0.59-1.47)
Cortical Atrophy *(yes vs. no)*					0.41 (0.11-1.37)	0.41 (0.11-1.39)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **NIA-RI** = National Institute of Aging – Reagan Institute; **OR** = Odds ratio

Table D14: The Association Between Grammatical Complexity and Cognitive Resilience Adjusted for Number of Cerebral Infarcts and Cortical Atrophy, Either CERAD or NIA-RI criteria (n=58)

Cognitive Resilience, Either CERAD or NIA-RI Criteria (n=58)

	Unadjusted Model	Age	*APOE*-ε4 Presence	Number of Infarcts	Cortical Atrophy	Full Model
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Variable						
Grammatical complexity *Low[1] vs. high*	**0.16 (0.03-0.59)**	**0.16 (0.03-0.60)**	**0.15 (0.03-0.56)**	**0.17 (0.04-0.63)**	**0.18 (0.04-0.67)**	**0.17 (0.03-0.67)**
Age (years)		1.00 (0.84-1.20)				1.00 (0.83-1.22)
APOE-ε4 Presence			0.37 (0.10-1.22)			0.38 (0.10-1.30)
Number of Infarcts				0.86 (0.55-1.23)		0.88 (0.56-1.30)
Cortical Atrophy *(yes vs. no)*					0.41 (0.12-1.31)	0.44 (0.12-1.48)

[1]Low was defined as the lowest quartile vs quartiles 2 to 4.

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **CI** = Confidence interval; **NIA-RI** = National Institute of Aging – Reagan Institute; **OR** = Odds ratio

Appendix E: Assessment of Selection Bias in Analytic Samples

To analyze the relationship between written language skills and cognitive resilience in the present study, many exclusions were made to create the analytic samples. Written language skills (idea density and grammatical complexity) both rely on the availability of autobiography data. In total, data from 180 autobiographies were available, meaning 498 individuals from the overall Nun Study (n=678) were excluded. Exclusions were also made from the analytic samples (for CERAD or NIA-RI criteria) if an individual did not meet the respective criteria. From the available autobiographies, 124 participants were excluded from the CERAD analytic sample (n=56). Similarly, from the available autobiographies, 138 participants were excluded from the NIA-RI analytic sample (n=42).

Given that many participants were excluded for not meeting the respective criteria, differences between the analytic sample and those who were excluded from participation were a concern. Assessments of these potential differences are provided in Tables E1 through to E3. In each table, the available population for age at death was smaller than the overall population or exclusion category. The smaller availability of data on age at death was due to some participants being alive at the time of data collection. Therefore, to better assess the variations in age between the groups, age at last cognitive assessment was also analyzed since all participants had data for this variable. Data on presence of *APOE*-ε4 were also not available for all participants, limiting some comparisons.

When the total Nun Study population was compared to the total autobiography sample, it was evident that those excluded from participation were significantly older (using both age at death ($p<.001$) and age at last cognitive assessment ($p<.001$) variables). The average age at death for excluded participants from the total autobiography sample was 91.24 in comparison to 87.87 in the total autobiography sample. However, excluded participants from both the CERAD (Table E2) and NIA-RI (Table E3) analytic samples were significantly younger. When comparing the CERAD sample to the excluded participants, the excluded participants were significantly younger at death (89.29 vs 87.04, $p<0.0007$) and at last cognitive assessment (88.66 vs 86.95, $p<.0001$) than those included in the CERAD analytic sample. Similarly, excluded participants from the NIA-RI analytic sample had significantly younger means for both age at death (89.19 vs 87.36, $p=0.009$) and age at last cognitive assessment (88.56 vs 87.15, $p=0.048$). The tendency for excluded participants from the analytic samples to be younger is not surprising. Older

participants who were included in both analytic samples would have lived longer and therefore would have been more likely to have developed the AD neuropathology that was required for inclusion.

Dementia status at last assessment also varied significantly for excluded participants from the total autobiography sample ($p=0.001$), the CERAD analytic sample ($p=0.0006$) and the NIA-RI analytic sample ($p<0.001$). The lower prevalence of dementia in the total autobiography sample may reflect the younger age of this sample. Given that the analytic sample was specifically limited to individuals who had Alzheimer neuropathology, a higher prevalence of dementia would be expected and this was found in the CERAD and NIA-RI analytic samples.

Significant differences were also found for level of education when comparing the total Nun Study to the total autobiography sample. Generally, excluded participants tended to have higher education than those included in the present study. This finding is not surprising since previous research has consistently linked higher educational attainment to longevity. In this study, individuals who were excluded (including living participants) would have been more highly educated. However, level of education did not differ significantly for either CERAD or NIA-RI when comparing excluded participants to the analytic samples. When comparing the presence of *APOE*, no significant differences were noted for the total autobiography sample (Table E1, $p=0.16$) or the CERAD analytic sample (Table E2, $p=0.19$). However, significant differences were noted for excluded participants from the NIA-RI sample, in terms of *APOE*-ε4 presence ($p=0.009$). Excluded participants from the NIA-RI sample tended to have more individuals without *APOE*-ε4 alleles present.

Comparisons of written language skills (among the total autobiography sample, the CERAD analytic sample, and the excluded participants) generally showed no significant differences for either idea density and grammatical complexity. However, idea density differed significantly ($p=0.04$) between excluded participants and the NIA-RI analytic sample.

The potential for selection bias in derivation of analytic samples must be considered. However, overall, where there were significant differences between samples, most were readily explained based on established associations with eligibility criteria, such as presence of Alzheimer neuropathology needed for the definition of the outcome of cognitive resilience; individuals meeting these criteria are a select sample.

Table E1. Comparison of Total Autobiographies Sample to Excluded Participants

Characteristic		Total Nun Study Population (n=678)	Total Autobiography Sample (n=180)	Excluded Participants from Total Sample (n=498)
Age at death, mean (SD)**		90.40 (5.37)[1]	87.87 (4.46)[2]	91.24 (5.39)[3]
Age at last cognitive assessment**		89.48 (5.71)	87.48 (5.26)	90.20 (5.70)
Level of education (%)*				
	<High school	10.03	5.00	11.85
	High school	5.46	3.89	6.02
	Bachelor's degree	39.82	38.33	40.36
	≥ Master's degree	44.69	52.78	41.77
Presence of *APOE*-ε4 (%)				
	No	77.22[4]	73.17[5]	78.68[6]
	Yes	22.78[4]	26.83[5]	21.32[6]
Dementia status at last assessment (%)**				
	No	56.19	66.67	52.41
	Yes	43.81	33.33	47.59

* p<.05; ** p<.01

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **NIA-RI** = National Institute of Aging – Reagan Institute; **SD** = Standard deviation

Note:

[1]For the total Nun Study population, n=606 had complete data for age at death

[2] For the total autobiography sample, n=151 had complete data for age at death

[3] For the excluded participants sample, n=455 had complete data for age at death

[4] For the total Nun Study sample, n=619 had complete data for presence of *APOE*-ε4

[5] For the total autobiography sample, n=164 had complete data for presence of *APOE*-ε4

[6] For the excluded participants sample, n=455 had complete data for presence of *APOE*-ε4

Table E2. Comparison of CERAD Sample to Excluded Participants Who Wrote an Autobiography

Characteristic	Total Autobiography Sample (n=180)	CERAD Analytic Sample (n=56)	Excluded Participants from CERAD Sample (n=124)
Age at death, mean (SD)**	87.87 (4.46)[1]	89.29 (3.06)	87.04 (4.94)[2]
Age of last cognitive assessment**	87.48 (5.26)	88.66 (3.16)	86.95 (5.91)
Level of education (%)			
<High school	5.00	3.57	5.65
High school	3.89	5.36	3.23
Bachelor's degree	38.33	42.86	36.29
≥ Master's degree	52.78	48.21	54.84
Presence of *APOE*-ε4 (%)			
No	73.17[3]	66.07	76.85[4]
Yes	26.83[3]	33.93	23.15[4]
Dementia status at last assessment (%)**			
No	66.67	48.21	75.00
Yes	33.33	51.79	25.00
Idea Density			
Lowest quartile (Q1)	25.00	23.21	25.81
Higher quartiles (Q's 2-4)	75.00	76.79	74.19
Grammatical Complexity			
Lowest quartile (Q1)	24.44	28.57	22.58
Higher quartiles (Q's 2-4)	75.56	71.43	77.42

* p<.05; ** p<.01

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **NIA-RI** = National Institute of Aging – Reagan Institute; **SD** = Standard deviation

Note:

[1] For the total autobiographies sample, n=151 had complete data for age at death

[2] For the excluded participants from CERAD sample, n=95 had complete data for age at death

[3] For the total autobiography sample, n=164 had complete data for presence of *APOE*-ε4

[4] For the excluded participants from CERAD sample, n=108 had complete data for presence of *APOE*-ε4

Table E3. Comparison of NIA-RI Sample to Excluded Participants Who Wrote an Autobiography

Characteristic	Total Autobiography Sample (n=180)	NIA-RI Analytic Sample (n=42)	Excluded Participants from NIA-RI Sample (n=138)
Age at death, mean (SD)**	87.87 (4.46)[1]	89.19 (3.39)	87.36 (4.73)[2]
Age at last cognitive assessment*	87.48 (76.67)	88.56 (3.33)	87.15 (5.69)
Level of education (%)			
<High school	5.00	4.76	5.07
High school	3.89	4.76	3.62
Bachelor's degree	38.33	45.24	36.23
≥ Master's degree	52.78	45.24	55.07
Presence of *APOE*-ε4 (%)**			
No	73.17[3]	57.14	78.69[4]
Yes	26.83[3]	42.86	21.31[4]
Dementia status at last assessment (%)**			
No	66.67	30.95	77.54
Yes	33.33	69.05	22.45
Idea Density*			
Lowest quartile (Q1)	25.00	38.10	21.01
Higher quartiles (Q's 2-4)	75.00	61.90	78.99
Grammatical Complexity			
Lowest quartile (Q1)	24.44	35.71	21.01
Higher quartiles (Q's 2-4)	75.56	64.29	78.99

* p<.05; ** p<.01

Abbreviations: ***APOE*-ε4** = Apolipoprotein E-ε4; **CERAD** = Consortium to Establish a Registry for Alzheimer's Disease; **NIA-RI** = National Institute of Aging – Reagan Institute; **SD** = Standard deviation

Note:

[1]For the total autobiographies sample, n=151 had complete data for age at death

[2]For the excluded participants from NIA-RI sample, n=109 had complete data for age at death

[3]For the total autobiography sample, n=164 had complete data for presence of *APOE*-ε4

[4]For the excluded participants from NIA-RI sample, n=122 had complete data for *APOE*-ε4