Multilingualism and the risk of Alzheimer disease and dementia

by

Erica Hack

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Science
in
Health Studies and Gerontology

Waterloo, Ontario, Canada, 2011

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

Background: Alzheimer disease (AD) is a progressive, late-life neurodegenerative disorder. Given the aging population, AD is a significant health concern. According to the Alzheimer Society of Canada (Smetanin et al., 2009), in 25 years 2.8% of the Canadian population will have AD or a related dementia. Presently, there is no cure for AD; therefore, efforts to either delay AD onset or prevent AD altogether are a primary focus.

The ability to proficiently speak many languages has been associated with certain cognitive advantages. Based on these findings, multilinguals are hypothesized to be more resistant to cognitive decline than monolinguals. More research is warranted in order to further this theory and to contribute to strategies to prevent or delay AD.

Objectives: The first study objective was to evaluate whether multilingualism was associated with the development of AD. The second study objective was to assess whether multilingualism was associated with later dementia onset.

Methods: Analyses were based on data from the Nun Study, a longitudinal study of aging in 678 participants 75+ years living in the United States. In order to address the first study objective, the association between multilingualism and AD was assessed in 157 participants using logistic regression models adjusted for age, education, apolipoprotein E-E4 (ApoE-E4) status, immigrant status, and occupation. Additional subgroup analyses also included covariates associated with career length and linguistic ability (grammatical complexity and idea density). AD was diagnosed based on criteria for both clinical dementia and AD neuropathology. Dementia was diagnosed using the Diagnostic and Statistical Manual of Mental Disorders, 4th edition criterion (American Psychiatric Association, 1994) (based on the Consortium to Establish a Registry for

Alzheimer's Disease battery of tests (Morris, Heyman, Mohs, & Hughes, 1989) and performance on activities of daily living), while AD neuropathology was based on the National Institute on Aging and Reagan Institute criterion (The National Institute on Aging - Reagan Institute (NIA-RI) Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease, 1997). In order to address the second study objective, dementia likelihood was assessed in 325 participants using discrete-time survival analyses adjusted for age, ApoE-E4 status, education, and linguistic ability.

Results: When adjusted for age, education, ApoE-E4 status, occupation, and immigrant status, participants speaking two or more languages had similar AD risks compared to monolinguals (OR = 1.05; 95% CI = 0.45-2.50). However, when grammatical complexity was held constant across participants, speaking two or more languages was associated with a four-fold decrease in AD risk compared to speaking one language (OR = 0.25; 95% CI = 0.04-1.23), although this did not reach statistical significance.

When the association between multilingualism and time of dementia onset was assessed, the dementia hazard function estimates for all participants were constant and persisted throughout the follow-up period of the study. When ApoE-E4 status and baseline age were held constant, participants speaking four or more languages were significantly less likely to develop dementia than monolingual participants (OR = 0.14; 95% CI = 0.01-0.66). An interaction between multilingualism and the other two covariates (ApoE-E4 status and baseline age) was observed: the oldest participants with an ApoE-E4 allele who spoke four or more languages had smaller dementia risks than younger participants without an ApoE-E4 allele who spoke one, two, or three languages. Participants speaking two or three languages were no less likely than monolinguals to develop dementia across the study duration. When idea density was held

constant across participants, multilingualism was associated with a nonsignificant decreased risk of dementia for individuals speaking three (OR = 0.62; 95% CI = 0.16-2.41) or four or more languages (OR = 0.53; 95% CI = 0.06-4.91) while participants speaking two languages were no more at risk for dementia than monolinguals (OR = 1.08; 95% CI = 0.43-2.69).

Discussion: Initially, multilingualism did not appear to confer protection against AD. After holding grammatical complexity constant across all participants, however, multilingualism was found to be associated with AD risk. Therefore, linguistic ability confounded the initial relationship measured by this study. When the association between multilingualism and time of dementia onset was evaluated, participants were no more likely to develop dementia in one time period than another, and monolingual participants were no more likely to develop dementia in earlier time periods than multilinguals. While a trend of decreasing dementia risk with ascending number of languages spoken was not observed, speaking four or more languages was consistently associated with decreased dementia risk compared to speaking one language. The presence of an ApoE-E4 allele and low linguistic ability had a strong and consistent significant association with increased AD and dementia risk. Therefore, the influence of these variables on the association of multilingualism with AD and dementia is worthy of further exploration.

Overall, this study provided some support for a protective effect of multilingualism on AD and dementia. Some of the present investigation's results differ, however, from those of previous studies. This is not surprising, considering the present study utilized different methodologies than other studies in this research area. For instance, our study employed a definition of multilingualism based on self-report data – participants were classified as multilingual based on the number of languages they reported proficiency with. Therefore, our

definition of multilingualism was less strict than definitions used in previous studies. However, our study employed much stricter outcome criteria than those used in previous studies, as our study is the first in this area to confirm AD cases with AD neuropathology evaluations. Our study is also the first io utilize prospective data and to include participants who remained dementia-free in addition to participants developing AD and dementia. In addition, this is the only study in this research area to evaluate the relationship of multilingualism with AD and dementia in the context of important covariates such as ApoE-E4 status and linguistic ability. Therefore, while some of our results contrast with other findings in this area, this is understandable given our novel methodologies. A broad range of study methods must be used in the future if we are to generate the depth of evidence needed for a full understanding of the relationship of multilingualism with AD and dementia. A better understanding of this relationship may also provide insight into both cognitive and brain reserve mechanisms, which could help more individuals maintain cognitive function into late life.

Acknowledgements

I would like to thank my supervisor, Dr. Suzanne Tyas, for her mentorship and support during the duration of my degree. I feel very privileged to have had a supervisor as attentive and generous as yourself, Suzanne. I feel as though your students and their learning are always your main priorities, despite your many other commitments. I truly appreciate all of the guidance you have given me throughout the past two years.

I would also like to thank my committee members, Dr. Joel Dubin and Dr. Myra Fernandes, for their help with the development of this thesis.

Lastly, to Mom, Dad, Michael, Richard, Courtney, Christine, and Karina: simply put, I could not have done this without you. I am truly blessed to have such great friends and a supportive and loving family.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Aβ plaques	Beta-amyloid plaques
$A\beta_{42}$	42 amino acid beta-amyloid polypeptide
ACh	Acetylcholine
AD	Alzheimer disease
ADL	Activity of daily living
ApoE-E4	ApolipoproteinE-E4 allele
APP	Amyloid precursor protein
CERAD	Consortium to Establish a Registry for Alzheimer's Disease
ChAT	Choline acetyltransferase
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, fourth edition
MMSE	Mini-Mental State Examination
NFT	Neurofibrillary tangles
NIA-RI	National Institute on Aging-Reagan Institute
NINCDS-ADRDA	National Institute of Neurological Disorders and Stroke-Alzheimer Disease and Related Disorders Association
NP	Neuritic plaque
OR	Odds ratio
SD	Standard deviation
SES	Socioeconomic status
95% CI	95 % Confidence interval

1.0 Introduction

Canada, similar to many other nations, has a large and growing population of citizens aged 65 and older. By the year 2026, when many citizens belonging to the "baby boom" generation will have retired, the median Canadian age is expected to be 43.3 years. At this time, citizens aged 65 and above will comprise 21.2% of the nation's population (Foot, 2008; Schellenberg & Turcotte, 2007). According to Health Canada (2001), 43% of all national health expenditures in 2000-01 were for the care of adults aged 65 and older. Thus, the aging population represents a pressing issue for our health care system and for the world alike.

Alzheimer disease (AD) is a progressive, late-life neurodegenerative disorder. Given the aging population, AD is a significant health concern. According to the 2010 Rising Tide report, the Alzheimer Society of Canada estimates that in 25 years, 2.8% of the Canadian population will have AD or a related dementia. By 2038, a new case of dementia is estimated to develop every 2 minutes (Smetanin et al. 2009). Currently, there is no cure for AD; thus, efforts to either delay AD onset or prevent AD altogether are a primary focus.

Research interest has recently shifted to the potentially neuroprotective influence of mental engagement, often by means of mentally stimulating activities. It has been proposed that the neural benefits of mental engagement are analogous to the physical benefits of regular aerobic exercise – one must regularly use an organ, otherwise they stand to "lose" it (Swaab et al., 2002). The more one is mentally active, the more reserve, or capacity for neurological insult, one is hypothesized to develop against neuropathology-induced cognitive decline (Stern, 2002). In light of this theory, it has been hypothesized that multilingualism, or the ability to speak more than one language, may help to protect against late-life cognitive decline. Given that multilingualism has been associated with other cognitive benefits, such as heightened

executive control and function, multilinguals might develop higher reserve (cognitive reserve, brain reserve, or both) levels than monolinguals over time. Recent findings suggest that this is indeed the case, as multilinguals are suggested to have later dementia onset compared to monolinguals (Craik, Bialystok, & Freedman, 2010; Bialystok, Craik, & Freedman 2007).

The first aim of the present study was to evaluate whether multilingualism was associated with an outcome of AD. The second aim of this investigation was to assess whether monolinguals were more likely to develop dementia earlier than their multilingual counterparts.

Our analyses were based on longitudinal cohort data from the Nun Study. In this study, multilingualism was ascertained by means of a self-report questionnaire. Participants were required to list the languages with which they had proficiency, in order to facilitate teaching placements overseas. The study also accounted for the possible effects of other factors, such as education, apolipoprotein E-E4 carrier status, and linguistic ability on the relationship of multilingualism with AD and dementia. Contributions from many types of studies assessing the association between multilingualism and late-life cognitive decline are needed in order to fully understand this research area.

Since a there is no cure for AD, it is imperative that potential protective factors are identified so that the future burden of AD may be diminished. A better understanding of multilingualism's relationship with AD and dementia can help to develop methods of AD protection and avoidance, as well as clarify reserve mechanisms. In turn, it is hoped that this information will help more individuals maintain cognition as they approach late life.

2.0 Literature Review

2.1 Alzheimer Disease

2.1.1 Epidemiology and Public Health Impact

Dementia is currently one of the most burdensome and devastating diseases facing society. Today there are more than 35 million people with dementia globally (Wimo & Prince, 2010) and about 4.6 million new cases are estimated to develop worldwide each year (Ferri et al., 2005). AD is the most common form of dementia and is the chief condition leading to nursing home placement (Farlow, 2010). According to the 2010 World Alzheimer Report released by Alzheimer's Disease International, an estimated $604 billion dollars US were spent in 2010 on direct health care, caregiving costs, and nursing home costs for dementia patients. These staggering costs account for approximately 1% of the world's gross domestic product, ranging from 0.24% of GDP from low income countries to 1.24% of GDP from high income countries (Wimo & Prince, 2010).

Currently 1 in 11 Canadians over age 65 have dementia and AD accounts for about 64% of these cases (Smetanin et al., 2009). In 2008, there were approximately 103 700 new cases of AD diagnosed in Canada, which equates to roughly one new case diagnosed every five minutes (Smetanin et al., 2009). The burden of this disease is likely to become even worse in the coming years. The reasons for this are twofold: AD incidence is expected to increase due to an increasing average age of the population, and the duration spent living with the disease is also expected to increase as a result of medical surveillance and treatment. In addition to the economic and health care burden, AD is also a source of emotional distress and burden on families; the Canadian Alzheimer's Society has reported the amount of time provided by family to care for AD patients to be 231 million hours in 2008 (Smetanin et al., 2009). Thus, it appears

that society would have much to gain, both economically and emotionally, if the effects of AD could be reduced or eliminated.

2.1.2 Alzheimer Disease Risk Factors

Age is the most obvious risk factor for AD. After age 65, the prevalence of AD is estimated to double approximately every five years (Jorm, Korten, & Henderson, 1987). The incidence of AD has been shown to increase from 1.5 per 1000 person-years in those aged 65-69 years to 52.6 per 1000 person-years in those aged 90+ years (Bermejo-Pareja, Benito-Leon, Vega, Medrano, & Roman, 2008).

Family history is another potent AD risk factor. In terms of genetic influences, AD is a heterogeneous disease that can be classified as either familial or sporadic. Familial AD is autosomal dominant and typically presents before age 65. Cases of familial AD are rare, with a prevalence below 0.1% (Harvey, Skelton-Robinson, & Rossor, 2003). Most cases of familial disease result from mutations in the presenilin 1 gene (chromosome 14) and presenilin 2 gene (chromosome 1); however, mutations in the amyloid precursor protein (APP) gene (chromosome 21) also contribute to a fraction of familial AD cases (Hardy, 1997). In terms of sporadic AD, the apolipoprotein E (ApoE) E4 allele accounts for most of the genetic risk (Raber, Huang, & Ashford, 2004). However, there may be other genetic influences on sporadic AD risk that are yet to be discovered. An apparent dose-response relationship has been observed between the number of E4 alleles, risk of AD development, and age of onset; when the number of E4 alleles increased from zero to two, AD risk increased from 20% to 90% and age of onset decreased from 84 to 68 years (Corder, Saunders, & Strittmater, 1993). It is hypothesized that ApoE facilitates amyloid deposition in the brain (Holtzman et al., 2000).

Besides age and genetics, the next most important AD risk factor is level of attained education. Numerous studies have reported an increased risk of AD among participants with lower levels of formal education (Gatz et al., 2001; Schmand et al., 1997; Stern et al., 1994). It is speculated that education may act, along with other "life course influential factors" (Richards & Sacker, 2003), such as occupation and early-life household socioeconomic status (SES), to modify other AD risk factors (e.g., brain size) and subsequent clinical manifestation (Borenstein et al., 2005; Karp et al., 2004; Moceri, Kukull, Emanuel, van Belle, & Larson, 2000; Moceri et al., 2001; Mortimer, Snowdon, & Markesbery, 2003; Stern et al., 1994). These risk factors are theorized to be connected through the concept of reserve (please refer to section 2.2).

Gender represents another potential AD risk factor. Female gender has been suggested to be a biologically plausible risk factor, as females appear to have greater age-related brain volume reductions than males, especially in the neural areas most affected in AD patients (Carr, Goate, Phil, & Morris, 1997). There has been some speculation that hormones, such as estrogen, may also lead to an increased AD risk in females (Janicki & Schupf, 2010). In spite of this logic, reviews and meta-analyses of gender-specific incidence studies have found female gender not to be significantly associated with AD risk (Gao, Hendrie, Hall, & Hui, 1998; Swanwick & Lawlor, 1999). Although it may thus be interpreted that gender is not as important a risk factor as to other potential covariates, evidence still exists to suggest that gender may be a significant effect modifier with respect to familial AD (Farrer et al., 1997; Launer et al., 1999). Furthermore, gender-specific AD incidence is difficult to determine and may be unreliable in populations of older adults as they usually contain fewer men than women (Fratiglioni et al., 1991; Jorm et al., 1987). Thus, it still remains to be resolved whether gender has a significant influence on AD development.

Several other potential AD risk factors exist that fall into the category of lifestyle factors. These risk factors may influence AD development by means of vascular mechanisms, and include hypertension, diabetes mellitus, and hyperlipidemia (see reviews by de la Torre, 2010; Breteler, 2000). As such, smoking has also been identified to be a risk factor for AD. A recent meta-analysis examining 14 non-tobacco industry-affiliated cohort studies revealed smokers to have a significantly increased AD risk compared to non-smokers (Cataldo, Prochaska, & Glantz, 2010). Smoking is also hypothesized to contribute to AD neuropathology through oxidative stress (see review by Markesbery, 1997; Tyas et al., 2003). Since smoking is related to several vascular factors, it can also be conceptualized as a vascular risk factor for AD. Other vascular factors, such as hypercholesterolemia and hypertension, have also been associated with AD (Breteler, 2000). It is hypothesized that these risk factors contribute to cerebral hypoperfusion (see review by de la Torre, 2010), resulting in clinical AD symptoms (see review by de la Torre, 2004; Skoog & Gustafson, 2006).

Conversely, alcohol has been hypothesized to reduce the risk of AD through vascular mechanisms, or by means of introducing antioxidants to the system, in the instance of wine (see review by Panza et al., 2008). Despite these proposed mechanisms and supporting evidence (Ruitenberg et al., 2002), overall findings regarding alcohol and AD risk are relatively inconclusive (see reviews by Panza et al., 2008; Tyas, 2001) and thus the role of alcohol intake in AD still remains to be clarified. Considering the multitude of suspected AD risk factors, in addition to the discrepancies between AD pathology and clinical manifestation (see Section 2.2), it is clear to see how AD etiology is particularly challenging to conceptualize.

2.1.3 Alzheimer Disease Etiology

Although the first case of AD was described over a century ago (Stelzmann, Schnitzlein, & Murtagh, 1995), there remains a substantial knowledge gap concerning the nature of AD pathogenesis. Given the many potential influences in its causal pathway, there exist many hypotheses concerning the origins of AD. It is unlikely, however, that these theories are exclusive; it is instead conceivable that the true underlying mechanism of AD consists of multiple proposed pathways interacting with one another.

In the 1970s the first AD etiological theory, known as the cholinergic hypothesis, emerged and began to gather support (Bartus, Dean, Beer, & Lippa, 1982; Cummings & Kaufer, 1996; Francis, Palmer, Snape, & Wilcock, 1999). This theory is rooted in the notion that AD is characterized by a decrease in acetylcholine (ACh) signal transmission in the central nervous system. The relationship first began to gain interest when the cholinergic system was observed to be related to cognitive dysfunction (Drachman & Leavitt, 1974). It was later confirmed that there was a cholinergic deficit associated with AD (Bowen et al., 1983; Coyle, Price, & DeLong, 1983; Sims et al., 1980), likely due to reduced ACh synthesis by the enzyme choline acetyltransferase (ChAT). Evidence of markedly reduced ChAT activity in Alzheimer brains compared with age-matched controls has lent support for this theory (Bird, Stranahan, Sumi, & Raskind, 1983; Nagai, McGeer, Peng, McGeer, & Dolman, 1983; Perry, Gibson, Blessed, Perry, & Tomlinson, 1977). Evidence of reduced ACh reuptake (Rylett, Ball, & Colhoun, 1983) and release (Nilsson, Nordberg, Hardy, Wester, & Winblad, 1986), in addition to loss of cholinergic neurons (Whitehouse et al., 1982) has also supported this theory. Since it is now clear other deficits and pathologies contribute to AD development, the original importance of the cholinergic hypothesis may have been overemphasized. Furthermore, contradictory

findings, such as ChAT up-regulation in some areas of the brain classically affected areas in AD (DeKosky et al., 2002), have weakened confidence in the cholinergic hypothesis. However, the cholinergic hypothesis was the main basis for the first AD pharmacological therapies and despite its downfalls and the advancements in current knowledge of AD, it is still the basis for most widely used AD pharmacotherapies.

Presently, the hallmark pathologic lesions of AD are intracellular neurofibrillary tangles (NFTs) and extracellular beta-amyloid (Aβ) plaques (also known as senile plaques, of which the etiologically important subtypes are neuritic plaques or NPs). Since dense-core amyloid plaques are thought to be specific to AD, whereas NFT have been observed in various neurodegenerative diseases, it has been argued that neural Aβ accumulation is key to AD pathogenesis (see review by Hardy & Selkoe, 2002). This supposition that Aβ plaques in neural tissue lead to the cognitive symptoms and neurodegeneration seen in AD is known as the amyloid hypothesis (see reviews by Hardy & Selkoe, 2002; Hardy, 2009; Hardy & Higgins, 1992), one of the leading theories concerning AD origins.

The plaques in question are deposits of Aβ peptides, cleavage products originating from a longer peptide known as amyloid precursor protein (APP). Initially it was thought that the Aβ found in plaques was an abnormal protein, but was later found to be a normal product of cell metabolism (Haass et al., 1992). APP, which has features of a transmembrane glycoprotein (Kang et al., 1987), is processed mainly by the non-amyloidogenic α-secretase within the Aβ sequence (Esch et al., 1990). In contrast, when APP is cleaved by the β-secretase and γ-secretase at its C-terminus and N-terminus, respectively, a 42 amino acid polypeptide is liberated from APP which is consequently secreted from cells (Estus et al., 1992; Golde, Estus, & Younkin, 1992; Haass et al., 1992). When the $A\beta_{42}$ product is released, it is then deposited in

the form of Aβ plaques (Borchelt et al., 1997; Citron, Oltersdorf, & Haass, 1992). Aβ plaques are thought to directly cause synaptic and neuronal injury, but also indirectly lead to inflammation via complement and cytokines, which may further aggravate neural damage (Akiyama et al., 2000). Some forms of familial early-onset AD, which represent a small proportion of all AD cases, are thought to result from an over-production of $A\beta_{42}$ peptides due to mutations in the APP genetic sequence (see review by Hardy & Higgins, 1992; St George-Hyslop et al., 1987). Increased $A\beta_{42}$ production is also seen in the other forms of familial early-onset AD (Scheuner et al., 1996; Tedde et al., 2003).

While the amyloid hypothesis offers a broad outline to explain AD pathogenesis and has led to clinically promising research, certain observations do not fit easily with its rationale. One of its strongest objections is that the number of neural amyloid plaques does not correlate well with the degree of observed neuronal death (Irizarry et al., 1997). For instance, Schmitz et al. (2004) found neuronal loss to be in hippocampal areas affected by Aβ deposition, but also in areas far from plaques. The degree of amyloid deposition also does not correlate with the degree of clinical impairment experienced in life. Many participants expressing neuropathological hallmarks of AD do not exhibit clinical symptoms of AD (Crystal et al., 1988; Katzman et al., 1988; Price & Morris, 1999; Snowdon et al., 1996). Furthermore, it has also been shown that plaque removal from cortical tissues does not ameliorate clinical symptoms or improve cognition (Holmes et al., 2008). These findings have created holes in the amyloid hypothesis and suggest involvement by more than one mechanism is likely at play with regard to AD pathogenesis.

NFTs, the other neuropathologic hallmark of AD, have also fuelled interest in terms of their etiologic role in AD. While Aβ plaques are extracellular deposits, NFTs are intracellular

aggregates of hyperphosphorylated tau protein (Grundke-Iqbal et al., 1986). Tau is a normal protein that binds to microtubules in order to promote their assembly and stability, and its phosphorylation is normally regulated by a balance between kinases and phosphatases (Iqbal et al., 2005). In AD, tau levels become higher than normal (Baas & Qiang, 2005) and the balance supposedly shifts towards tau hyperphosphorylation in an attempt to deal with these elevated levels. The hyperphosphorylated tau then clumps together in aggregates called NFTs. Tau hyperphosphorylation also leads to microtubule disassembly and compromised neuronal and synaptic function, since normal tau and other microtubule-associated proteins are also affected when tau is hyperphosphorylated (Iqbal et al., 2005). It was initially unclear whether treatments directed at AD tau pathology would prevent Aβ-induced impairments, since it had been suggested that Aβ plaques precede or initiate NFT formation (Busciglio, Lorenzo, Yeh, & Yankner, 1995; review by Hardy & Selkoe, 2002). However, this notion has not been supported by any substantial evidence and, conversely, it has been argued that NFT formation precedes amyloid deposition and plays a more central role to AD pathogenesis (Braak & Braak, 1991). Tau might simply modulate the effects of Aβ plaques and their toxicities (Rapoport, Dawson, Binder, Vitek, & Ferreira, 2002; Roberson et al., 2007). Furthermore, NFT quantities have also been shown to better predict cognitive function, compared to Aβ plaques (Giannakopoulos et al., 2003). Currently, the relationship between the two neuropathologic hallmarks is still largely unknown.

There is mounting evidence to suggest that, in sporadic AD, Aβ plaques and NFTs are merely downstream indicators of damage resulting from preceding vascular factors. The vascular hypothesis of AD broadly asserts that cardiovascular pathology results in brain hypoperfusion, which in turn is the principal biological instigator of cognitive decline (see

review by de la Torre, 2010). This theory is consistent with the finding that cerebrovascular disease, such as stroke and atherosclerosis, is present in a large percentage of AD patients (see reviews by de la Torre, 2004; Kalaria, 2003). Cerebral capillary degeneration has also been shown to exist in essentially all AD brains examined post-mortem (de la Torre, 2002). Furthermore, ApoE-E4 is believed to be a strong risk factor for cerebral amyloid angiopathy, which predisposes individuals for cerebral haemorrhages and endothelial dysfunction (see review by Kalaria, 2003). According to a 2004 review by de la Torre, there is already an extensive amount of epidemiologic, pharmacologic, and neuroimaging-based evidence to support the vascular hypothesis of AD and the volume of supportive findings continues to grow.

Other hypotheses exist concerning the pathogenesis of AD; however, they are less established and may tie in with the previously outlined hypotheses to help explain existing knowledge gaps or discrepancies. For instance, age-related myelin breakdown has been proposed as a primary disease process in AD (see review by Bartzokis, 2004). This is due to reports of myelin breakdown occurring in AD participants during the earliest, or even preclinical, stages of AD (Bartzokis et al., 2003; review by Bartzokis et al., 2004; de la Monte, 1989). It has also been argued that demyelination is accelerated or initiated by amyloid deposition (Kurt et al., 2001).

Another potential contributor to AD pathogenesis is inflammation. Animal models and clinical studies strongly suggest that inflammation significantly contributes to AD pathogenesis (see reviews by Akiyama et al., 2000; Wyss-Coray, 2006); whether inflammation is a principal instigator or secondarily exacerbates neural damage has yet to be determined. Similarly, AD has been associated with large amounts of oxidative stress (see review by Butterfield &

Lauderback, 2002), but it is also unknown whether this is a primary or secondary contributor to the disease process (see review by Markesbery, 1997).

The disjunction between hallmark AD pathologies and clinical expression, as well as the other gaps in current etiological theories, has led to the development of the concept of reserve. This will be discussed in later sections (see Section 2.2).

2.1.4 Alzheimer Disease Diagnosis

Although diagnoses of "probable" AD may be given to individuals presenting with symptoms during life, a "definite" diagnosis of AD is not assigned until after death, when a post-mortem brain autopsy can be performed. Therefore, there are two aspects to a AD diagnosis: one conducted during life (the neuropsychological examination) and one conducted after death (the neuropathological examination). A definitive diagnosis of AD is made only when both the neuropsychological (clinical) and neuropathological data are consistent with AD.

Neuropsychological evaluation is critical for establishing the nature of cognitive impairment and the extent of behavioural disability. Presently, several different criteria for AD exist based on neuropsychological examinations. The three most frequently used criteria for making clinical AD diagnoses are the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) (Morris, Heyman, Mohs, & Hughes, 1989), the National Institute of Neurological Disorders and Stroke-Alzheimer Disease and Related Disorders Association (NINCDS-ADRDA) (McKhann et al., 1984) and the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) (American Psychiatric Association, 1994). Each criterion is comprised of similar general components. For instance, each criterion requires a comprehensive clinical interview, including patient history of AD symptoms, and physical and neurological examinations to rule out other systemic or brain

diseases that may account for cognitive deficits. Each criterion also has its own battery of neuropsychological tests. These tests assess several aspects of cognition: memory, orientation, language/verbal fluency, praxis, attention, and problem solving. Most diagnoses of clinical AD require deficits in two or more of these cognitive areas; these deficits must be severe to the point of impairing activities of daily living (ADLs). Any deficits in memory or other areas of cognition must also be insidious over the course of 6-12 months, depending on the criterion employed. At the present time there are no criteria that require the use of imaging techniques, such as PET or MRI scanning; however, these tools are often employed in order to gather additional neuroanatomical information before a diagnosis is given. Neuroanatomical features characteristic of AD include atrophy of the cerebral cortex and ventricular enlargement (see review by Blennow, de Leon, & Zetterberg, 2006). Furthermore, AD is also associated with degeneration in the medial temporal lobe, including the hippocampus and entorhinal cortex (see review by Braak et al., 1999). Depending on availability, other tools, such as biomarker evaluation, may also be used to make a diagnosis. The employment of such tools is not currently required by any clinical AD criteria; however, many AD researchers and clinicians currently believe these diagnostic guidelines should be updated in order for these modern techniques to be maximally utilized (see reviews by Dubois et al., 2007; Kolata, 2010; The National Institute on Aging - Reagan Institute (NIA-RI) Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease, 1997). Although a less popular idea, it has also been proposed that the future DSM-V definition of AD be further subdivided into categories according to secondary behavioural characteristics (Jeste, Meeks, Kim, & Zubenko, 2006) since it is arguable that these characteristics are more distressing, costly, and impairing than the cognitive symptoms of AD. Regardless of the criteria used to make a clinical

diagnosis of AD, the diagnosis remains presumptive until it is confirmed by neuropathologic examination.

Like the neuropsychological criteria, several post-mortem neuropathological definitions of AD exist. There are three that are most commonly used and recognized by current neuropathologists: Braak's staging method (Braak & Braak, 1991), the criterion of the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) (Mirra et al., 1991), and the National Institute for Aging, Ronald and Nancy Reagan Institute of Alzheimer's Disease criterion (NIA-RI) (The National Institute on Aging - Reagan Institute (NIA-RI) Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease, 1997). Generally, AD neuropathology classification is based on the quantification and distribution of the AD neuropathologic hallmarks: NPs and NFTs. The CERAD criterion assigns diagnoses based principally on information about NP densities and locations, using the diagnostic categories 0 (no NPs), A, B, and C (frequent NPs) (Mirra et al., 1991; The National Institute on Aging - Reagan Institute (NIA-RI) Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease, 1997). Braak's staging method, however, assigns disease progression stage based solely on NFT distribution and location using the categories 0-VI, with stage VI relating to frequent NFTs in critical neural areas (Braak & Braak, 1991). The NIA-RI method examines both markers of AD neuropathology, as it measures NP densities as well as uses Braak staging to quantify NFTs (The National Institute on Aging - Reagan Institute (NIA-RI) Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease, 1997). The diagnostic categories employed by the NIA-RI criterion range from "low likelihood" to "high likelihood" of clinical dementia resulting from AD neuropathology. Both the CERAD and NIA-RI criteria incorporate

clinical information into the neuropathological diagnosis, or likelihood that the individual in question had dementia due to AD. All of the mentioned AD neuropathologic guidelines have their own strengths and weaknesses. For instance, while the CERAD criterion allows the consideration of other non-AD disorders, such as Parkinson's disease and Lewy body dementia (Mirra et al., 1991), it also bases AD diagnoses on solely NPs, which do not correlate with the degree of cognitive impairment. To base a diagnosis on solely NFTs (using Braak staging), the former issue is eliminated; however, quantities of NFTs are known to increase with age and are non-specific to AD (see review by Braak et al., 2003). Moreover, they are also found in cognitively normal elders (Knopman et al., 2003). The NIA-RI may at first appear to be an optimal criterion since it considers both NPs and NFTs, but many brains fall into diagnostic categories that cannot be defined by the NIA-RI diagnostic criterion (Nelson, Kukull, & Frosch, 2010). Therefore, it is apparent why controversy still exists when designating one criterion as the "best" diagnostic method for AD neuropathology.

2.2 Cognitive Reserve

2.2.1 Cognitive Reserve Definition

The evidence concerning the etiological role of NPs and NFTs in AD is indisputably convincing. However, since discrepancies between the degree of brain pathology and clinical expression still exist, the hypothetical construct of reserve has been proposed (see review by Stern, 2002) (Figure 1). Reserve can be conceptualized in two different ways: brain reserve and cognitive reserve. Brain reserve, a passive reserve model concerning the "hardware" of the brain, speculates that quantitative, structural entities, such as brain size and number of synapses, confer a certain capacity to endure neuropathologic insult. Thus, a brain with more neurons

would be able to remain above the "threshold" of disease manifestation despite having the same amount of neuropathology as another brain with less brain reserve. Cognitive reserve, alternatively, is an active reserve model concerning the "software" of the brain and its ability to recruit neural networks, or to use them more efficiently, in order to optimize cognitive status or maintain cognition despite the existence of pathology. A review by Scarmeas et al. described cognitive reserve with an analogy: "a trained mathematician...might be able to solve a mathematics problem many different ways, while a less experienced individual might have only one possible solution strategy available" (2003, page 631). Given that the brain reserve model alone cannot explain differences in clinical manifestation despite identical pathology, and that many factors associated with cognitive reserve also impact brain reserve and vice-versa, it is likely that brain reserve and cognitive reserve are implicated in the clinical manifestation of AD.

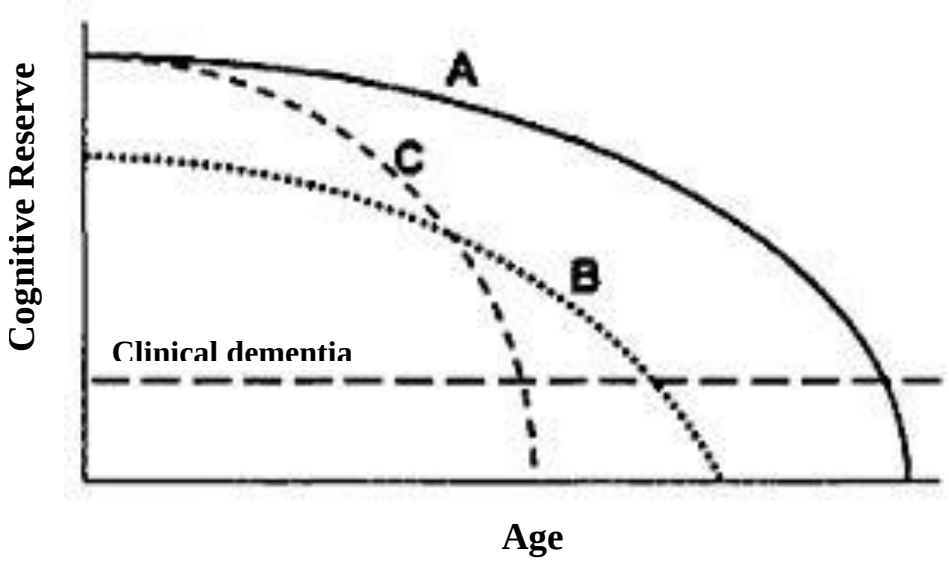

Figure 1. Adapted from Graves (2004).Theoretical diagram of cognitive reserve and how it may reconcile individual differences in clinical expression despite similar levels of AD neuropathology. Line "A" represents the trajectory for individuals in the population who do not develop AD or signs of clinical dementia in their lifetime. Line "B" represents the trajectory for individuals who have the same rate of AD neuropathology development as "A" but have lower levels of cognitive reserve and thus develop clinical dementia. Line "C" represents the trajectory of individuals with similar levels of cognitive reserve relative to individuals in "A" but have a genetic predisposition to AD.

The brain reserve and cognitive reserve constructs relate to the notion of AD being a malleable "life course" disease (Richards & Sacker, 2003; review by Richards & Dreary, 2005) – that is, disease contributors and protective factors potentially intervene across the entirety of one's lifespan, leading to late-life disease manifestation. While cognitive functions (such as memory, processing speed and verbal abilities) have been shown to have genetic foundations (see review by Lee, 2003), manifestation is differential depending on the influences (both positive and negative) during a given lifespan. Each risk or protective factor is thought to contribute independently to reserve, although the weight each carries in determining reserve levels is still unknown (Richards & Sacker, 2003; Stern et al., 1994). This implies that both brain reserve and cognitive reserve are dynamic entities, resulting from many exposures or combinations thereof, and thus can be modified during one's lifetime. Many AD risk factors implicated in altering one's level of brain reserve are those contributing to a reduction in neural structures or brain size (see reviews by Stern, 2002; Stern, 2009). Examples of such factors include early-life environmental factors (e.g., childhood nutrition), toxin exposure (e.g., lead), and head injury. Cognitive reserve, on the other hand, is traditionally related to factors associated with life experiences. While this is generally true, some factors may influence both cognitive and brain reserve, which suggests that both types of reserve are ultimately intertwined and can impact one another with respect to maintaining cognition in late life. For instance, Maguire et al. (2000) studied the hippocampal volumes of British taxi drivers and found posterior hippocampal volumes to positively correlate with career length. Given the preferential involvement of this neural area when learned spatial information is recalled, this example illustrates how experiential factors may also have implications for brain reserve. Another example of experience-mediated structural changes was provided by Elbert, Pantev, Wienbruch,

Rockstroh, & Taub (1995), who found that professional string musicians had larger cortical representations corresponding to the fingers they used most during string instrument performances. According to reviews by Daffner (2010) and Stern (2006), transgenic AD mice placed in intellectually enriching environments (i.e., those with more objects or toys for the mice to explore) have shown reduced AD neuropathology levels, suggesting mental stimulation may influence neural structures and protect against AD beyond the scope of cognitive reserve. These findings imply that brain reserve is not entirely dependent on factors with biological bases (such as genetics). It is important to know how cognition is affected by multilingualism (i.e., whether neural structures [brain reserve], their efficiencies [cognitive reserve], or both are influenced by multilingualism), but overall it is more clinically relevant to clarify whether multilingualism is associated with improved outcomes concerning late-life cognitive decline, regardless of the mechanism.

2.2.2 Component Factors

As indicated by Stern (2009), cognitive reserve can be conceptualized as two components: neural reserve, which relates to the individual differences in healthy brain network efficiencies; and neural compensation, which relates to the individual differences in the network's ability to compensate for neural pathology. Since cognitive reserve concerns how brain structures are utilized to process tasks and not necessarily the structures' sizes or densities, suggested contributors of cognitive reserve are those that create higher levels of mental engagement (also known as "intellectual stimulation" or "cognitive stimulation") (see review by Daffner, 2010; Wilson, Barnes, & Bennett, 2003). According to a 2002 review by Stern, the most commonly studied cognitive reserve measure is level of educational attainment. While it can be argued that education is a measure of innate intelligence, which is a passive reserve

correlate, other evidence suggests educational attainment influences cognitive reserve above and beyond innate intelligence (Evans et al., 1993). When using education as a CR measure, it is important to note potential pitfalls, such as when participants have lower educational attainment due to unrelated external influences (e.g., limited educational opportunities). Nevertheless, education is usually a relatively easy variable to ascertain and also can serve as a proxy for other cognitive reserve measures, such as IQ.

High levels of formal education are more likely to encourage intellectual challenges, which are hypothesized to spur the development of more complex and efficient neural connections. Many past evaluations of education's influence on late-life cognitive outcomes were conducted with populations where higher education was relatively uncommon, either due to limited opportunities or events such as war or famine. Thus, it will be interesting to compare past results with future findings based on current populations, considering that obtaining a university education is more common in the present day.

Occupation is another frequently considered variable when evaluating cognitive reserve. Generally, occupations of higher rankings (Richards & Sacker, 2003) or requiring complex skills (Le Carret et al., 2003) have been associated with higher cognitive reserve. Since higher-level occupations usually require advanced educations, occupation and formal education may initially appear to be interchangeable. However, occupation is thought to add to the effect of education (Evans et al., 1993; Richards & Sacker, 2003; Stern et al., 1994; Stern, Tang, Denaro, & Mayeux, 1995; Le Carret et al., 2003) as it has been found that participants with professional occupations maintain or increase their cognitive abilities with their occupational practice. These findings suggest that the continual maintenance, or challenge, of intellect could be just as important to cognitive reserve as the attainment of a given educational level. Occupation might

even prove to be a superior measure of continued cognitive challenge over the lifespan, rather than a mere measure of educational attainment, given the length of time one potentially spends in their career versus obtaining an education.

According to a 2010 review by Daffner, physically and cognitively stimulating activities may be the most influential agents (external to biological and genetic factors) when determining successful cognitive aging. Indeed, it has been shown that a higher level of lifetime engagement in mentally stimulating leisure or social activities is predictive of higher cognitive performance, and thus higher cognitive reserve, in late life (Fratiglioni, Paillard-Borg, & Winblad, 2004; Staff, Murray, Deary, & Whalley, 2004; review by Valenzuela & Sachdev, 2006). Examples of studied activities include playing games such as chess or cards, reading books, and participation in social activities like volunteering (see review by Scarmeas & Stern, 2004; Solé-Padullés et al., 2009; review by Stern, 2006; Wilson, Barnes, & Bennett, 2003). According to reviews by Daffner (2010) and Stern (2006), intellectually enriching environments have also been shown to reduce the accumulation of AD neuropathologies in transgenic AD mice, suggesting mental stimulation may protect against AD onset beyond just the scope of cognitive reserve. These measures are less frequently used in the literature than education and occupation; however, this is most likely due to limited data concerning these factors. Furthermore, some activities are applicable to only certain communities or cultures; thus it may be difficult to measure only certain activities if the population is culturally diverse. A particular strength of using mentally stimulating activities as cognitive reserve measures is that these measures differ from educational and occupational measures since they are more likely to be dissociated from SES, which is a common confounder (Wilson et al., 2003). On the other hand, a potential limitation to measuring mental stimulation is that the degree of challenge required by a given activity may

be entirely dependent on the context. For instance, an activity that seems challenging to one person may not be particularly challenging to another, due to the diverse cognitive ranges found within a given population. The degree of mental challenge experienced would also depend on learning style, previous experience with similar challenges, and a multitude of other factors that appear implausible to measure for experimental purposes.

In addition to education, occupation, and participation in mentally stimulating activities, a high level of linguistic ability has been also linked to the avoidance of cognitive decline (Snowdon et al., 1996; Solé-Padullés et al., 2009; review by Valenzuela & Sachdev, 2005). In the past, linguistic ability has been most explicitly researched in the Nun Study, which is unique in its ability to assess linguistic ability through participants' handwritten autobiographies (Snowdon et al., 1996). Linguistic ability has also been shown to be a strong predictor of asymptomatic AD (Tyas, Snowdon, Desrosiers, Riley, & Markesbery, 2009), which is when cognition is maintained despite the presence of post-mortem AD neuropathology. This evidence implies that linguistic ability might serve as a cognitive reserve correlate. In the Nun Study, linguistic ability was classified into the sub-variables grammatical complexity, which differs according to sentence structure and forms of embedding, and idea density, or the average number of ideas expressed per ten words (Snowdon et al., 1996). Increased overall literacy was also found to be associated with a slower decline in memory (Manly, Schupf, Tang, & Stern, 2005; review by Stern, 2009): in two studies of cognitively normal elders, Manly et al. (2003 and 2005) found participants with lower literacy had steeper delayed recall score declines over time (three years and five years, respectively) compared to the more highly literate participants. These results suggest higher literacy skills may slow age-related memory decline. Furthermore, regression models from the 2005 study indicated those with lower literacy levels had

significantly higher rates of executive function and language decline compared with highly literate participants (Manly et al., 2005).

According to reviews by Scarmeas et al. (2003) and Stern (2009), linguistic ability is related to educational attainment; however, linguistic ability has yet to be further expanded to include other potentially relevant sub-variables that may not be as strongly connected to formal education. For instance, a given level of linguistic ability might also depend on the degree of participation in activities involving reading and writing, social connectivity, leisure activity participation, or complexity of career-related pursuits. The relationship may additionally work in the reverse: possessing a high level of linguistic ability could predispose, or motivate, one to participate in more of the previously specified activities.

Multilingualism, or the ability to proficiently speak more than one language, is another variable related to linguistic ability that may factor into cognitive reserve capacity. Proficiently speaking multiple languages, as well as switching between languages, can be viewed as a mentally challenging process. Additionally, in some cases this mode of mental stimulation is utilized every day, which could equate to countless hours of perpetual stimulation. Therefore, it is conceivable that proficiently speaking more than one language may enhance cognitive reserve. In support of this hypothesis, Chertkow, Whitehead, Wolfson, Atherton, & Bergman (2010) and Bialystok et al. (2007) have demonstrated multilinguals to have lower rates of cognitive decline when compared to their monolingual counterparts. These relationships were significant regardless of level of formal education. Further findings from studies in this area will be discussed in Section 2.3. Since the data on these variables have been limited thus far, it still remains to be seen how multilingualism, as well as all potential linguistic ability variables, may relate to cognitive reserve capacity. While speaking multiple languages may assist mental acuity

(by developing higher attentional control, for instance – this will be developed further in Section 2.3), so little is known about cognitive reserve and influential factors in cognitive reserve development that other influential factors relating to cognitive reserve should be considered to be equally important. For instance, the knowledge of many languages would also broaden the number of social interactions available to an individual, which subsequently could lead to the engagement in unique recreational, or occupation-related, activities. A resulting higher level of engagement in these forms of activities also may influence cognitive reserve. More conclusive information needs to be known first about these different factors before the different factors can be valued in relation to each other with respect to the outcome of cognitive reserve enhancement or maintenance.

2.2.3 Implications and Outcomes in Alzheimer Disease Research

There are many points along the natural history of AD that are of interest with regard to alleviating the burden of the disease. As such, a variety of outcomes can be measured in an attempt to detect the existence and mechanism of cognitive reserve. In the most basic sense, cognitive reserve can be assessed by comparing those who develop AD and those who do not. In the past, most studies have done this by comparing incident AD, usually only in the clinical sense ("probable AD" according to the NINCDS-ADRDA criterion (McKhann et al., 1984); another correlate would be clinical dementia). Those who do not manifest clinical AD despite having substantial AD neuropathology are usually believed to have a higher cognitive reserve than those who do. This outcome requires longitudinal cohort study data, which are collected using an expensive and elaborate prospective cohort design. Despite these drawbacks, several cohort studies have found evidence to support relationships between AD risk and some cognitive reserve factors, such as education (Borenstein et al., 2005; Karp et al., 2004; Stern et

al., 1994), occupation (Qiu et al., 2003; Stern et al., 1994) and intellectually stimulating activities (Wilson et al., 2003). Studies assessing incident AD with neuropathologic confirmation are less common than clinical AD as post-mortem information is needed in order to ascertain this outcome.

Another outcome of interest based on solely clinical data is delay of AD onset. According to the 2010 Rising Tide report, even a two-year delay in dementia onset over the next ten years (and subsequent decrease in AD prevalence) has the potential to reduce the total economic burden in Canada by $24.2 billion dollars (Smetanin et al., 2009). Therefore, investigating whether certain exposures induce a delay of disease onset is relevant to the public health burden. Participants manifesting AD symptoms at older ages than other participants are hypothesized to do so because of having higher cognitive reserve levels. It is also hypothesized that participants with higher cognitive reserve decline faster once AD has manifested compared to those with lower cognitive reserve (see review by Scarmeas & Stern, 2003).

Despite existing epidemiological evidence, many questions still exist with regard to cognitive reserve and AD manifestation, especially in terms of individual differences in life experiences (see review by Scarmeas & Stern, 2004). For instance, very little is known about how life experience variables interact with genetic risk factors, such as ApoE-E4 status. As mentioned previously, the possession of at least one ApoE-E4 allele has been shown to strongly impact the development of sporadic AD; however, this influence may be altered depending on the spectrum of intervening life experience variables. Considering it is one of the most important biological risk factors for AD, ApoE-E4 status is likely to confound or modify the relationship between life experience variables and AD (Borenstein et al., 2005). Furthermore, since a majority of the research about cognitive reserve factors and AD have utilized solely

clinical data, more studies of the relationship between cognitive reserve variables and AD as confirmed by neuropathology would also add to the growing body of cognitive reserve knowledge. Cognitive reserve is an appealing concept as it makes logical sense and suggests that our neuropsychological fates are not necessarily sealed by structural entities alone. However, more in-depth research assessing associations between variables implicated in cognitive reserve and structural, or biological, variables must first be conducted before the concept is embraced and applied clinically.

2.3 Multilingualism and Cognitive Reserve

2.3.1 Multilingualism and Cognition in Older Adults

As previously outlined, linguistic ability may serve as an indicator of mental engagement and the ability to recruit different cognitive pathways as one ages. In light of this evidence, other variables related to linguistic ability have been investigated with respect to whether they provide any cognitive advantage. Since the process of switching between two languages is considered to be cognitively demanding, it has been suggested that multilingualism, or the ability to proficiently speak more than one language, may promote a higher level of mental engagement and thus enhance cognitive reserve. Furthermore, the ability to speak multiple languages could also enhance mental stimulation by means of the increased likelihood of socialization, or participation in diverse activities. At the moment, the relationship between multilingualism and AD has been probed using preliminary investigations. If a positive association can be established between multilingualism and higher late-life cognitive ability (or cognitive reserve) we may gain insight into the etiology of late-life neurocognitive impairments, such as AD. Likewise, this information may also aid in future efforts addressing the prevention of late-life cognitive decline, or the maintenance of healthy cognitive states in advanced age.

Investigations comparing cognitive differences between multilinguals and monolinguals have largely concerned two main cognitive areas: verbal fluency, which concerns the ability to generate words based on pictures or belonging to a certain category (Rosselli et al., 2000), and executive control, which broadly encompasses task planning and the ability to organize behaviours so that one may self-monitor actions, ignore distracters, and be cognitively flexible (Appendix A, Table 1; see Appendix B for a summary of the various cognitive tests used by these studies of multilingualism and cognition). Many kinds of assessment are used to evaluate

executive control; examples of such include the Simon task (Bialystok, Craik, Klein, & Viswanathan, 2004; Fernandes, Craik, Bialystok, & Kreuger, 2007), saccadic eye movements (Bialystok, Craik, & Ryan, 2006), and the Stroop test (Bialystok, Craik, & Luk, 2008). All methods are used to measure the ability to correctly perform a task while ignoring irrelevant stimuli, also known as divided attention, which is closely related to working memory (Bialystok et al., 2006; Bialystok et al., 2008; Craik & Bialystok, 2006).

With respect to multilingualism and potential cognitive advantages, verbal fluency and executive control have been of most interest since they appear to reflect differing inherent strengths possessed by individuals depending on their status as a monolingual or multilingual (Bialystok et al., 2008). Multilinguals have demonstrated cognitive advantages over monolinguals in the area of executive control, whereas they have shown cognitive disadvantages compared to monolinguals in the area of verbal fluency.

2.3.1.1 Multilingualism and Executive Control

Since multilinguals must resist speaking an alternate language while speaking another, they exercise continual resistance from interfering stimuli. Thus, multilinguals have been hypothesized to have advantages over monolinguals with respect to executive control (Bialystok et al., 2008). Comparisons of executive control in older monolinguals and older multilinguals have supported this theory by demonstrating evidence of a modest, yet significant, multilingual advantage in executive control (Appendix A, Table 2). An example of such evidence was provided by Bialystok et al. (2004). In this study, the Simon Task was employed to measure executive control. The Simon Task requires participants to respond correctly to a given stimulus; the challenging aspect of this test is that the stimulus does not logically match, or is incongruent to, the desired response. When the investigators compared the response times of

younger participants (monolinguals and multilinguals) to that of older participants (monolinguals and multilinguals), older participants had slower responses (an expected finding). However, the older multilinguals appeared to exhibit significantly less of an age-related slowing of response time when compared to older monolinguals, once adjusted for age, education, and SES. Therefore, multilingualism was proposed to be the quality that allowed these participants to process and give faster responses. In 2008, Bialystok et al. replicated the study and these results; the older monolinguals again exhibited significantly slower responses to incongruent tasks while the older multilinguals and both younger groups' responses were unaffected by the change in task congruency. The authors also employed the Stroop task (where participants name the colour of a word rather than reading the word itself) to measure executive control; however, no significant differences between groups were found using this task. This was unexpected, as it was hypothesized that advantages in executive control should be demonstrated by multilinguals regardless of the task used for executive control assessment.

Another evaluation of attentional control in older monolingual and older multilingual adults provided some interesting results concerning potential multilingual advantages in this area of cognition. Fernandes et al. (2007) found that when older monolinguals and multilinguals were asked to recall words either with or without attentional distracters present during encoding (viewing the words and committing them to memory) or retrieval (recalling the previously viewed words) processes, older monolinguals recalled a greater number of words, on average, than older multilinguals in the full attention condition (without distracters). Recall in bilinguals, however, was similar to that of monolinguals in the majority of divided attention conditions. Although the investigators failed to find a conclusive multilingual advantage in resisting the effects of interference, their results suggest that multilingualism may have had beneficial effects

in reducing the disadvantage seen in the full attention condition. While these findings are somewhat inconsistent with other evidence showing clearer multilingual advantages in attentional control, the results of Fernandes and colleagues could have been attributable to the established association between multilingualism and smaller vocabulary size, as well as the documented multilingual disadvantages with lexical access (see Section 2.1.3.2). For instance, multilinguals could have experienced difficulties in verbal recall if they had smaller overall vocabularies than monolinguals. Additionally, these analyses did not adjust for many covariates, meaning these differences could have been due to factors other than differences in language fluency.

It is currently unclear whether confounding variables or interactions have significant roles in this observed association, due to the developing nature of the research area. In 2006, Bialystok et al. found results similar to their 2004 study by measuring executive control using saccades rather than performances on the Simon Task. There appeared to be a significant interaction between age and language group, as older multilinguals were not significantly slower in responding than their younger counterparts, whereas older monolinguals were slower than young monolinguals. This study, however, did not appear to account for other potentially confounding variables such as education or SES; thus, these results must be considered with some caution as interactions or confounding from these variables could have been present.

It is also unclear how an advantage in executive control could relate to advantages in overall cognition. Craik and Bialystok (2006) measured executive control differences between multilinguals and monolinguals using a task-simulation test, which was a simulation of a task characteristic of daily life. This task required participants to plan and execute "cooking breakfast". Ultimately, this evaluation method tested aspects otherwise not captured by other

kinds of executive control assessments, as it was able to functionally test prospective memory and working memory in addition to executive control: participants had to "set the table" in the midst of the "cooking breakfast" task. Younger and older groups of both monolinguals and multilinguals were also compared in this investigation. Although the investigators found no interaction between age group and language fluency status, older multilinguals did spend significantly less time inappropriately "setting the table" (setting the table when they should have been focused on cooking the meal) than older monolinguals. No advantage was found for either monolinguals or multilinguals in terms of working or prospective memory. This suggests that, despite similar capacities for some aspects of executive function, older multilinguals appear to maintain an advantage over older monolinguals when ignoring irrelevant stimuli. Whether, or how, this advantage relates to overall age-related cognitive functioning is still relatively contentious.

2.3.1.2 Multilingualism and Verbal Fluency

In contrast to disadvantages in executive control compared to multilinguals, monolinguals are thought to have verbal fluency advantages over multilinguals. Since monolinguals exclusively use one language, rather than divide their time between two or more, they are hypothesized to develop more complex vocabularies (Bialystok et al., 2008; Rosselli et al., 2000). For example, Rosselli et al. (2000) assessed verbal fluency among older monolinguals and multilinguals using the Boston Naming Test (BNT). A standard test used in verbal fluency assessments, the BNT scores participants based on their ability to name as many of objects as possible in pictures presented to them. Participants in this study were also asked to generate as many words as possible from phonemic categories (e.g., words beginning with "F") and semantic categories (e.g., animals), each in one-minute intervals. This investigation found that

older monolinguals produced significantly more words in semantic categories than older multilinguals. This finding held true regardless of whether the testing language was the multilinguals' first or second language. This difference did not remain significant, however, when the numbers of words generated in phonemic categories were compared. Despite not adjusting for any typical confounding factors (e.g., education), these results still provide a basis for hypothesis generation.

More recently, a 2008 study by Bialystok et al. tested differences in verbal fluency between older multilinguals and monolinguals using three tests: the Peabody Picture Test, an assessment where the participant must match an object to the word the experimenter specifies; the BNT; and fluency tests (semantic and phonemic). After adjustment for age, it was found that monolinguals significantly outperformed multilinguals using all three forms of assessment. This supports the result found by Rosselli et al. (2000) and the theory of monolinguals having verbal fluency advantages over multilinguals. Interestingly, when the results from older participants were compared to their younger, university-aged counterparts, the older adults were found to have generated more words. This finding also supports the hypothesis that, due to a longer time period for vocabulary accumulation, older participants are advantaged in the area of verbal fluency as compared to younger participants (see review by Burke & Shafto, 2008).

Since this research area is still in its infancy, many questions remain as to how, or if, the cognitive differences between multilinguals and monolinguals relate to age-related cognitive advantages. It is also unclear whether these cognitive advantages would differ between the normative and pathological aging processes. This is even more puzzling considering both verbal fluency and executive control are included in the standard cognitive testing batteries but are not necessarily valued differently in clinical settings. Moreover, many other cognitive domains

(e.g., recall, orientation) factor into the neuropsychological batteries used for clinical dementia assessments. In a 2005 review, Craik and Bialystok proposed that executive control advantages are related to overall cognitive processing abilities, whereas verbal fluency advantages are related to enhanced levels of raw vocabulary knowledge. Since both of these cognitive areas are thought to change differentially with age (Craik & Bialystok, 2005), it is still unclear how advantages in either may confer functional benefit. Research investigating cognitive differences between multilinguals and monolinguals in the context of AD may help to provide some insights for this research area: preliminary evidence in this area is discussed in the following sub-section.

2.3.2 Multilingualism, Cognitive Decline, and Alzheimer Disease

Relatively few studies of cognitive decline have focused on multilingualism as the exposure of interest (Appendix A, Table 3). This is likely due to the developing nature of the research area, in addition to a lack of data concerning principal variables. Both Bialystok et al. (2007) and Chertkow et al. (2010) studied the association between multilingualism and age at clinical AD diagnosis, as assessed by the NINCDS-ADRDA neuropsychological criterion. However, the two studies differed in multilingualism classification, since Bialystok et al. (2007) dichotomously classified participants as bilingual or monolingual while Chertkow et al. (2010) classified participants according to number of languages spoken (monolinguals vs. bilinguals vs. those speaking three or more languages for the majority of their lifespans). Bialystok et al. (2007) found monolinguals to be significantly younger at dementia diagnosis when compared to those proficiently speaking more than one language (Appendix A, Table 4). Chertkow et al. (2010) found that the age at dementia diagnosis was significantly higher in only those who spoke three or more languages, while the age at dementia diagnosis among the monolinguals and bilinguals

was essentially equivalent. This finding is puzzling, as it suggests that proficiency in two languages confers no cognitive benefit over proficiency in only one language, and contradicts Bialystok et al.'s research concerning advantages associated with bilingual advantages. While this result opposed the authors' initial hypothesis, they did not offer a theory as to why this result occurred. They did note, however, that their analyses did not account for potential confounders such as immigration status, genetics, and SES, and that bilingual advantages (i.e., advantages present in those fluent in two languages versus one) may only emerge only if these other factors are comparable across groups. After diagnosis, the rates of cognitive decline between language groups did not appear to be significantly different in either study. These results favour the idea that multilingualism enhances cognitive reserve, as one would expect participants with higher reserve capacities to have later onset of AD but then have similar rates of decline to those who experienced AD onset earlier. Both analyses, however, also found that formal education did not significantly influence the results, which was surprising since education is considered to be an influential covariate with respect to AD risk and cognitive reserve.

In 2010, Craik et al. published results of an investigation that was essentially a reproduction of Bialystok et al.'s 2007 study. Different participants were selected from the same memory clinic to evaluate the same research question used in the 2007 study. This 2010 study found results similar to that of Bialystok et al. (2007), which could be interpreted as support for their hypothesis that bilinguals have later AD onset. This follow-up study did not use any novel covariates or analyses compared to the original study; the follow-up study was also subject to all of the same potential biases present in the 2007 study. Both of these studies did not consider participants who were cognitively normal or did not develop AD as all participants were AD

cases from memory clinics, so the proportion of monolinguals and multilinguals in the population not developing AD was unknown.

A 2008 study (Kavé, Eyal, Shorek, & Cohen-Mansfield, 2008) also examined the relationship between multilingualism and cognitive decline. Although the authors did not examine AD specifically as an outcome of interest, they did base outcomes on scores from cognitive tests similar to those used in clinical AD assessments. These tests included the Katzman cognitive screening test (Katzman et al., 1983), which is used for assessing orientation, immediate and delayed recall, and concentration; and the Folstein Mini-Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975), which assesses orientation, immediate and delayed recall, attention, calculation, and executive function. Similar to Chertkow et al. (2010), Kavé et al. (2008) classified participants as speaking one language, two languages, or three or more languages. Kavé et al. found the number of languages spoken to be significantly associated with cognitive test performance; a greater number of languages spoken was associated with higher cognitive test scores across three "waves" of neuropsychological testing. This significant relationship persisted despite adjusting for years of formal education. Age and gender were also significant covariates, whereas birthplace and age of immigration were not significant.

Research suggesting that multilingual participants are older at AD onset supports the theory that multilingualism enhances, or is associated with, cognitive reserve. However, this research area still contains many knowledge gaps. Therefore, much more needs to be known about cognitive reserve, how it relates to AD, and how multilingualism relates to both before any definitive conclusions can be made. In order to address these gaps, investigations concerning these relationships should be diversified in order to gain new insights. For instance, each of the three studies evaluating the relationship between multilingualism and AD recruited participants

from memory clinics located in large Canadian cities (Bialystok et al., 2007; Craik et al., 2010; Chertkow et al., 2010). When participants are recruited from memory clinics, not all incident cases are captured – only cases presenting to clinics are recruited. Furthermore, if a study population consists entirely of participants with the outcome (as in the case of AD patients recruited from clinics), no information concerning exposed or unexposed members of the population without the outcome can be analysed. Participants presenting to memory clinics may also be more socially connected than undetected AD cases, as AD cases are more likely to present to medical aid if family or friends are concerned for their health (Connell & Gallant, 1996). This is an important consideration since social connectivity is hypothesized to enhance cognitive reserve, and could be a confounding influence. It is also possible that participants recruited from these sites may have been systematically different than cases from the general Canadian population. For instance, participants recruited from these urban clinics might be systematically different than patients from more remote or rural areas of Canada. AD patients living in rural areas also may not have the same access to specialist care as city-dwelling patients. This is particularly relevant given that a principal barrier in seeking medical attention for suspected AD has been shown to be limited access to a specialist (Connell & Gallant, 1996; Connell, Roberts, McLaughlin, & Carpenter, 2009).

Another limitation associated with the use of Canadian populations for analyzing this research question is that results may be confounded by SES (Morton & Harper, 2007; Morton & Harper, 2009; Colzato et al. 2008). It has been suggested that Canadian populations, especially those from Canada's larger cities which typically have larger proportions of immigrants, may not be best suited to the study of multilingualism and cognition, given that SES is large determinant of health outcomes. Furthermore, it has also been suggested these populations may be especially

difficult to analyze given that most immigrants to Canada are highly educated yet may still have low SES levels.

This research area also has a great deal of potential with regard to possible study outcomes. Past outcomes in this research area have been based on clinical information alone. A definite diagnosis of AD, however, requires additional information regarding the extent of AD neuropathology. This is needed so that cases of clinical dementia resulting from alternative causes (e.g., Lewy body dementia, frontotemporal dementia) can be separated from the true AD cases. Using a heterogeneous dementia outcome is problematic, as different dementia subtypes result from varying etiologic pathways and have different risk factors. While research using only clinical information is certainly still of value and the most practical option for most researchers, investigations employing information on AD neuropathology in addition to clinical data could also help form theories about potential interactions between life experience-related exposures (such as multilingualism) and biological factors, such as genetics. Therefore, a possible avenue for future research would be to examine the association between multilingualism and AD in the context of data on both clinical dementia and AD neuropathology.

Age at AD onset is undoubtedly an important outcome of interest; however, it is important to acknowledge that this outcome can be subject to certain errors and confounders. For example, the recorded age at AD diagnosis can depend on other unrelated factors, such as frequency of physician visits, referral wait times, and SES. Evaluating the roles of covariates and confounders in the context of the relationship between multilingualism and AD represents an large area with research potential. Few studies concerning multilingualism and cognition consider the effects of formal education, SES, and occupation: these classical confounders and would be especially relevant to studies concerning cognitive reserve. Additionally, ApoE-E4 status has

never been considered as a study covariate, which is important given that the presence of an ApoE-E4 allele has been found to strongly increase the likelihood of sporadic AD. Other potential covariates for consideration include alcohol use, smoking status, and gender.

Cognitive reserve outcomes also extend much further than a delay in AD onset or AD symptom expression. Therefore, investigating the relationship between multilingualism and AD incidence also would provide conceptual insights with regard to cognitive reserve. In order to measure such a relationship, a cohort study of older adults without AD would need to be followed over time, until AD developed in some participants.

Finally, novel analytic methods stand to reveal unique findings with respect to this relationship. Evidence concerning this research area has been based primarily on cross-sectional data. Cross-sectional studies, while cost-effective and useful for hypothesis generation, are limited in terms of potential analyses as they ascertain both exposure and outcome at the same point in time. As such, studies of this design cannot ascertain incident cases, and therefore cannot make speculations about AD risk. Cross-sectional studies are also limited in terms of establishing temporality between exposure and outcome, and inferring subsequent causality. In the case of multilingualism and AD, it is unlikely that reverse causality occurs between exposure and outcome in any study design; however, the establishment of temporality is nonetheless important. Longitudinal cohort studies, in contrast, are more suitable for testing previously established hypotheses and allow for outcome development in a previously established outcome-free population. These observations illustrate the methodological gaps in the literature and represent potential future avenues for this developing research area.

2.4 Summary

Due to its relevance to the cognitive reserve model, as well as the promising results of preliminary investigations, multilingualism appears to be a budding exposure of interest with regard to future epidemiological assessments of AD. Currently, evidence concerning the association between multilingualism and AD is limited, as are the methods used to assess such a relationship. For example, none of the previous studies in this area have confirmed cases of AD with post-mortem AD neuropathology. Furthermore, very few covariates or common confounders have been considered in the relationship between multilingualism and AD. In order to formulate more concrete theories regarding the relationship between multilingualism and AD, it is essential that investigations use a diverse range of study designs and analyses are adjusted for relevant covariates. Thus, the aim of the proposed investigation is to critically evaluate the association between multilingualism and neuropathologically confirmed AD, as well as dementia. This was done by considering both clinical and neuropathologic data from a longitudinal cohort and adjusting for relevant covariates.

3.0 Study Rationale and Research Questions

3.1 Study Rationale

The purpose of this study was to investigate the relationship of multilingualism with AD and dementia. This objective was meant to advance the understanding of how exercises of mental stimulation, particularly those relating to cognitive reserve, may potentially protect against AD and dementia. As this research area is still developing, there were several existing knowledge gaps. By using prospective data from a cohort study, this investigation had the opportunity to address some of these gaps and provide a unique perspective on the relationship in question.

This investigation considered information from all developing cases of AD and dementia in the study population. Thus, selection biases were minimized as cases were not differentially selected from the population. Furthermore, all study participants had similar incomes, medical access, and social connections, which minimized issues concerning income or social inequalities within the sample. This issue was particularly relevant to this research area, as other investigations have been criticized for not adequately controlling for SES disparities between language proficiency groups (Morton & Harper, 2007; Morton & Harper, 2009).

Past examinations of the association between AD and multilingualism have also had restricted capacities for covariate evaluation. For example, the influence of ApoE-E4 status has never been assessed. Other important covariates include education, linguistic ability, occupational status, immigrant status, and SES. Since all of these factors have been significantly associated with sporadic AD risk, an investigation considering these variables was warranted. The proposed investigation had genetic and occupational information on participants, and all participants were comparable according to adult SES. Thus, these covariates were accounted for

in this study. This investigation also controlled for many other possible confounders, such as access to medical care, tobacco use, and gender, as study participants were all female with similar lifestyle habits and social interactions.

Measurement of AD risk cannot be calculated using cross-sectional data, as incidence cannot be determined using a cross-sectional design. Therefore, past investigations in the area of interest have not provided evidence regarding AD or dementia risk reduction. The use of longitudinal data allowed our study to ascertain incident cases, which made these calculations possible. Changes in cognitive status over time were also monitored using these data, as cognitive follow-ups were conducted annually over 12 waves. This permitted the comparison of times to dementia development between different language proficiency groups.

3.2 Research Questions

The aim of the present investigation was to address the following two research questions:

Research Question 1:

Does multilingualism reduce the risk of developing AD?

Research Question 2:

Is multilingualism associated with a later onset of dementia?

4.0 Methods

4.1 Literature Search

With regard to the multilingualism section in the literature review, two literature searches were conducted in October 2010 in order to evaluate the existing research in this area. The first search was conducted using the Medline database (1950 to present). This search used the terms "multilingualism" and "aging" to search all fields. All fields were searched, instead of using only Medical Subject Heading (MeSH) terms, in order to conduct the most exhaustive search of relevant studies in this research area. This search was restricted to articles using human participants, articles written in English, and articles including participants in the "aged" category (65 years and older). This search initially yielded 22 articles and two books. The titles and abstracts of these publications were then read and were excluded if they i) did not use aged (65+ years) participants; ii) did not compare cognitive function between groups of monolingual and multilingual participants; or iii) concerned only the validation of multilingual versions of cognitive tests. After this exclusion, eight articles remained for appraisal.

A second literature search was conducted to supplement the previous search using the PsycINFO database (1840 to present). This search used the descriptor terms ("bilingualism" or "multilingualism") and ("aging" or "Alzheimer's disease" or "late life"). This search was restricted to articles written in English and using human participants and initially yielded 32 journal articles. The same exclusion criteria used in the previous search were applied to these articles. After exclusion in this manner, seven articles remained; however, six of these articles overlapped with those found by the earlier Medline search. Thus, one article found using the PsycINFO search was retained for appraisal. Additionally, one recently published article (Craik et al., 2010) was also included in the review as it was directly related to the research area and

had not been assigned search terms at the time the searches were performed. Once this article and the articles found in PsycINFO were combined with those found in the Medline search, a total of ten articles were retained for appraisal in this section of the literature review (Bialystok et al., 2008; Bialystok et al., 2007; Bialystok et al., 2006; Bialystok et al., 2004; Chertkow et al., 2010; Craik et al., 2010; Craik et al., 2006; Fernandes et al., 2007; Kavé et al., 2008; Rosselli et al., 2000).

The cognitive reserve section of the literature review was meant to provide background for the proposed investigation. Therefore, this search was not as exhaustive as that performed for the multilingual section. For this section, two searches were conducted in October 2010 – one searching the Medline database using PubMed (1950 to present) and one searching the PsycINFO database (1887 to present). Results were limited to those in English and to journal articles. The Medline search used the following search terms: ((("Cognition/physiology*"[MAJR] OR "Memory/physiology*"[MAJR] OR "Brain/physiology*"[MAJR] OR "Mental processes/physiology*"[MAJR] OR "Recruitment, Neuropsychological/physiology*"[MAJR] OR "Cognition disorders/physiopathology"[MeSH] OR "Neuronal plasticity/physiology*"[MAJR]) AND ("Alzheimer disease/physiopathology"[MESH] OR "Alzheimer disease/pathology"[MESH])) AND "cognitive reserve"[tiab]. "Cognitive reserve" was added to the search in the title and abstract since there is no current MeSH/Major term for this area. This search yielded 15 initial results. The PsycINFO search employed the descriptor terms ("disease course*" or "cognitive processes*") and ("Alzheimer's disease*" or "aging*") and the term "cognitive reserve" in the abstract. This was done since there is no formal search term relating to cognitive reserve in the PsycINFO database.

After the elimination of duplicate results, the abstracts of the initial search results were then subjected to exclusion criteria. Articles were excluded if they did not relate to the etiology of neurodegenerative disorders or if cognitive reserve was not discussed in terms of "cognitive experience" or "life experience" exposures. Review articles were retained as several reviews composed by cognitive reserve pioneers (e.g., Y. Stern) have provided valuable insights to the concept. Using these exclusion criteria, the 15 articles found using Medline were pared down to seven results, while the 17 articles found using PsycINFO were pared down to ten articles. Additionally, relevant articles cited by review papers found in the literature search were also used in the literature review in order to provide a comprehensive background on the research topic.

4.2 Data Source: the Nun Study

4.2.1 Study Population

The Nun Study is a longitudinal study of aging with a principal focus on investigating AD etiology and risk factors. The study originally began in 1986 as a pilot study on aging using data collected from the School Sisters of Notre Dame religious congregation in Minnesota. In 1990, the study was expanded to include participants living in other regions of the United States. Participant recruitment occurred during 1991-1993; all members of the School Sisters of Notre Dame born before 1917 were invited to join the study. Of 1 031 eligible participants aged 75 years and above at baseline, 678 (66%) agreed to participate in all aspects of the study. This included consent to a review of medical and archival records, annual cognitive and physical assessments, and brain donation after death. Study participants and non-participants did not differ significantly by mean age, country of birth, annual mortality rate, or race. From the Nun

Study cohort of 678, more than 90% have died; therefore, nearly the entire cohort has been followed to completion. Of the participants who have died, 547 brains have been received; most of which (to date) have had a completed neuropathologic evaluation. Few studies in the world have a larger set of donated brains with supplemental clinical data collected during life.

One of the many advantages of using the Nun Study data for this investigation is that participants were exposed to relatively similar lifestyle and environmental risk factors during their adult lives, which minimizes confounding from such variables. All participants were members of the School Sisters of Notre Dame religious congregation and had similar incomes, SES, social activities, social supports, marital and reproductive histories, and tobacco and alcohol use. Most participants had similar occupations, as most were teachers; however, some participants did occupy domestic jobs for various reasons. With regard to medical access, all participants had equal access to health services throughout their adult lives.

4.2.2 Data Collection

Cognitive and physical assessment data were collected at routine annual intervals after study enrolment. This investigation had access to 12 waves (including baseline) of cognitive and physical assessment data. Cognitive function was assessed by trained gerontologists using the CERAD battery of neuropsychological tests (Morris et al., 1989). Global cognitive functioning was assessed by scores from the MMSE (Folstein, et al., 1975). Functional ability in activities of daily living (ADLs), which include everyday tasks such as bathing and feeding oneself, were assessed by performance-based tests. For more details on these tests please see Tyas et al. (2007), Kuriansky & Gurland, (1976), and Potvin et al., (1972).

Assessments of AD pathology (i.e., NPs and NFTs) were conducted by a single blinded neuropathologist who classified neuropathology in a consistent fashion across all participants.

Data on participant ApoE genotype were collected using standard methods (Saunders et al., 1996) from buccal cells from living participants or brain tissue obtained at time of autopsy. Individuals that conducted genotyping were blinded to participant cognitive status.

Data retrieved from convent archives included information on participant education, language proficiency, linguistic ability, usual occupation, and number of years spent as a teacher (if the participant was a teacher). Information concerning these variables was originally ascertained by means of self-report. Language proficiency was reported by means of a self-report questionnaire that had been administered by the convent in 1983. Data on linguistic ability were obtained by assessing autobiographical essays written by participants in early adulthood (18 to 32 years of age; mean 22 years). Out of the total Nun Study population, 180 participants provided autobiographical essays and thus provided data on these variables. The methods by which these variables were coded are described in more detail in the following description of study covariates (Section 4.4.3).

4.3 Thesis Project

4.3.1 Research Question 1

4.3.1.1 Analytic Sample Derivation

A primary analytic sample composed of 157 deceased Nun Study participants who had complete information on the exposure, outcomes, and particular covariates of interest was used for the main Research Question 1 analyses. Figure 2 outlines how the analytic sample was derived from the original Nun Study population. Participants missing information on one or more variables of interest were excluded from the analytic sample. Participants included in the analytic sample were also required to either meet the criterion for a control or the criterion for

an AD case; any participants not meeting one of these two criteria were also excluded from the analytic sample. Appendix C (Table 1) provides an in-depth description of excluded observations.

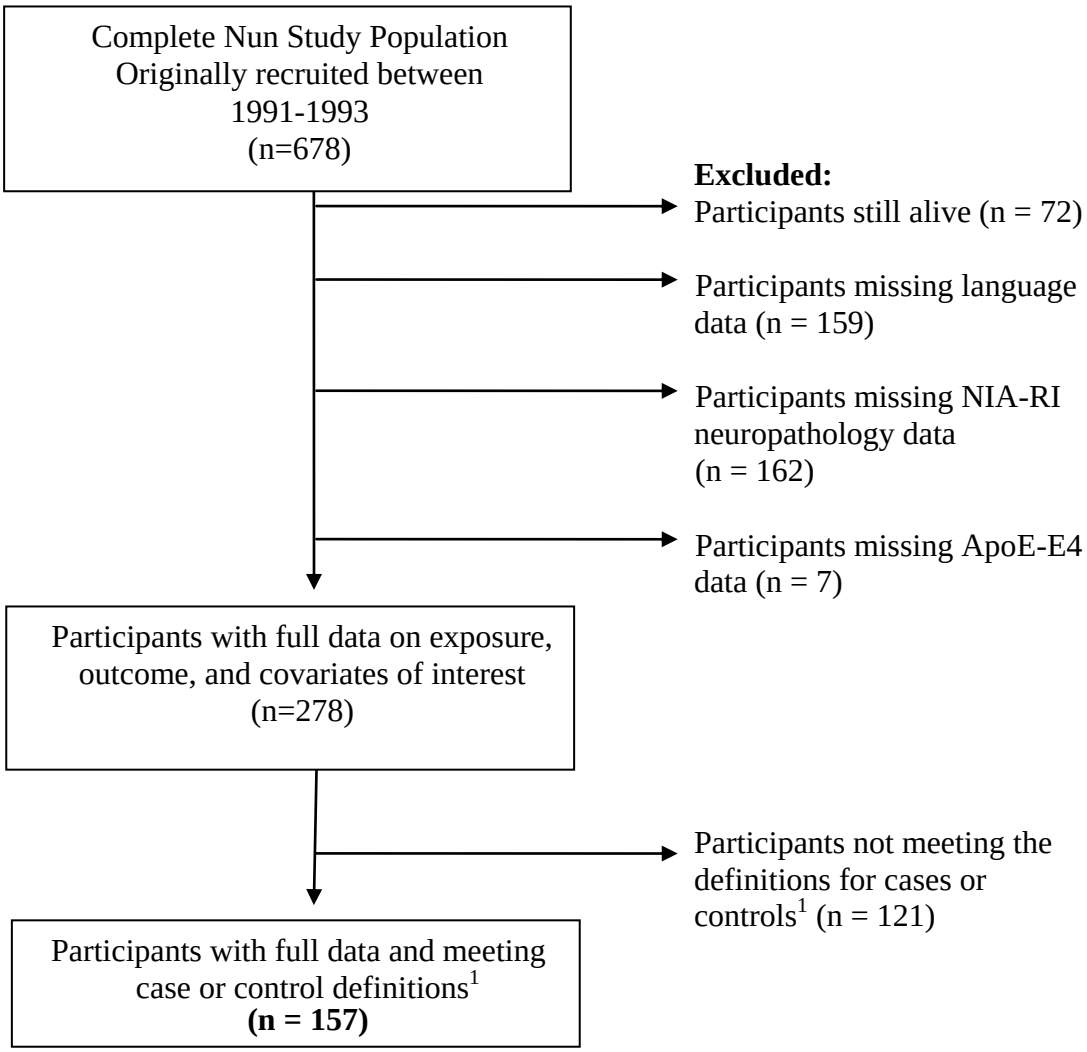

Figure 2. Derivation of the analytic sample used to address the main Research Question 1 analyses. Please see Appendix C (Table 1) for a summary of differences in characteristics between the excluded participants and those remaining in the analytic sample.
[1] Case = participant was clinically demented and exhibited "high likelihood" AD (The National Institute on Aging - Reagan Institute (NIA-RI) Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease, 1997). Control = participant was not clinically demented and exhibited "low likelihood" AD (The National Institute on Aging - Reagan Institute (NIA-RI) Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease, 1997). NIA-RI = National Institute for Aging, Ronald and Nancy Reagan Institute of Alzheimer's Disease; AD = Alzheimer disease; ApoE-E4 = Apolipoprotein E-E4

As many participants were excluded for not meeting the case or control criteria (n=121), significant discrepancies between analytic sample participant characteristics and characteristics of participants excluded from the analyses (non-response biases) were a concern. However, most aspects of this analytic sample were statistically similar to those of the excluded participants (see Appendix C, Table 1); one exception was that participants in the analytic sample were significantly older at last cognitive assessment (mean = 90.2 years; SD = 5.0 years) than excluded participants (mean age = 89.3 years; SD = 5.9 years; $p<0.05$). Thus, participants included in the analytic sample may have been more likely to live longer than participants excluded, as they may have been able to attend cognitive assessments longer into the study follow-up period than excluded participants. Despite this difference, the use of strict case and control criteria ensured our participants definitively had AD (cases) or were without AD (controls). Participants included in the analytic sample were not significantly different from excluded participants in terms of linguistic ability indicators.

4.3.1.2 Sensitivity Analysis Sub-Sample Derivation

In order to consider the impact of linguistic ability on the association between multilingualism and AD, a sensitivity analysis was conducted. This sensitivity analysis utilized the sample of participants from the original analytic sample with data on linguistic ability (Figure 3). Participants retained in this sub-sample had complete data on multilingualism, AD, original covariates of interest, and variables concerning linguistic ability (i.e., grammatical complexity and idea density). Two participants were also excluded from this analysis, as they were identified to be more influential than the other participants using standard diagnostic residual plots and criteria (see Section 5.2.1). Comprehensive descriptions of these influential participants and the process of their exclusion are included in Appendix H.

The characteristics of participants included in this sub-sample were compared to the characteristics of participants excluded from the sub-sample (as they were missing data on linguistic ability) but were included in the original analytic sample (Appendix C, Table 3). Participants missing linguistic ability data were significantly older at last cognitive assessment than the participants having linguistic ability data (91.2 years (SD = 5.1) versus 87.7 years (SD=3.7); $p<0.001$). A significantly greater proportion of participants missing linguistic ability data developed AD, compared to participants who had data on linguistic ability (47.7% versus 26.1%; $p<0.05$). Therefore, many participants with AD were not captured by this analysis and may not represent the general group of participants who had outcomes of AD. This sub-sample was much smaller (n=46) than the primary analytic sample used in Research Question 1 (n=157). Only 180 participants from the total Nun Study population provided handwritten autobiographies from which linguistic ability data were derived; therefore, the potential sample was limited even before additional participants were excluded for other missing data.

In order to consider the impact of career length on the association between multilingualism and AD, a second sensitivity analysis was performed in a sub-sample of only those participants from the original analytic sample who had occupied teaching positions (Figure 4). Since the majority of Nun Study participants were teachers, restricting this sub-sample to teachers did not exclude many participants (n=11 excluded; sub-sample size of n=146). A large proportion (90.9%) of participants excluded from this sub-sample had attained only grade school level education, which significantly differed from the educational levels of participants included in the sub-sample (0.7% of these participants had attained only grade school-level education, while 95.8% had university-level education). The excluded participants

also were older at last cognitive assessment than participants retained in this sensitivity analysis (93.4 years (SD=5.8) versus 89.9 years (SD=4.9); $p<0.05$).

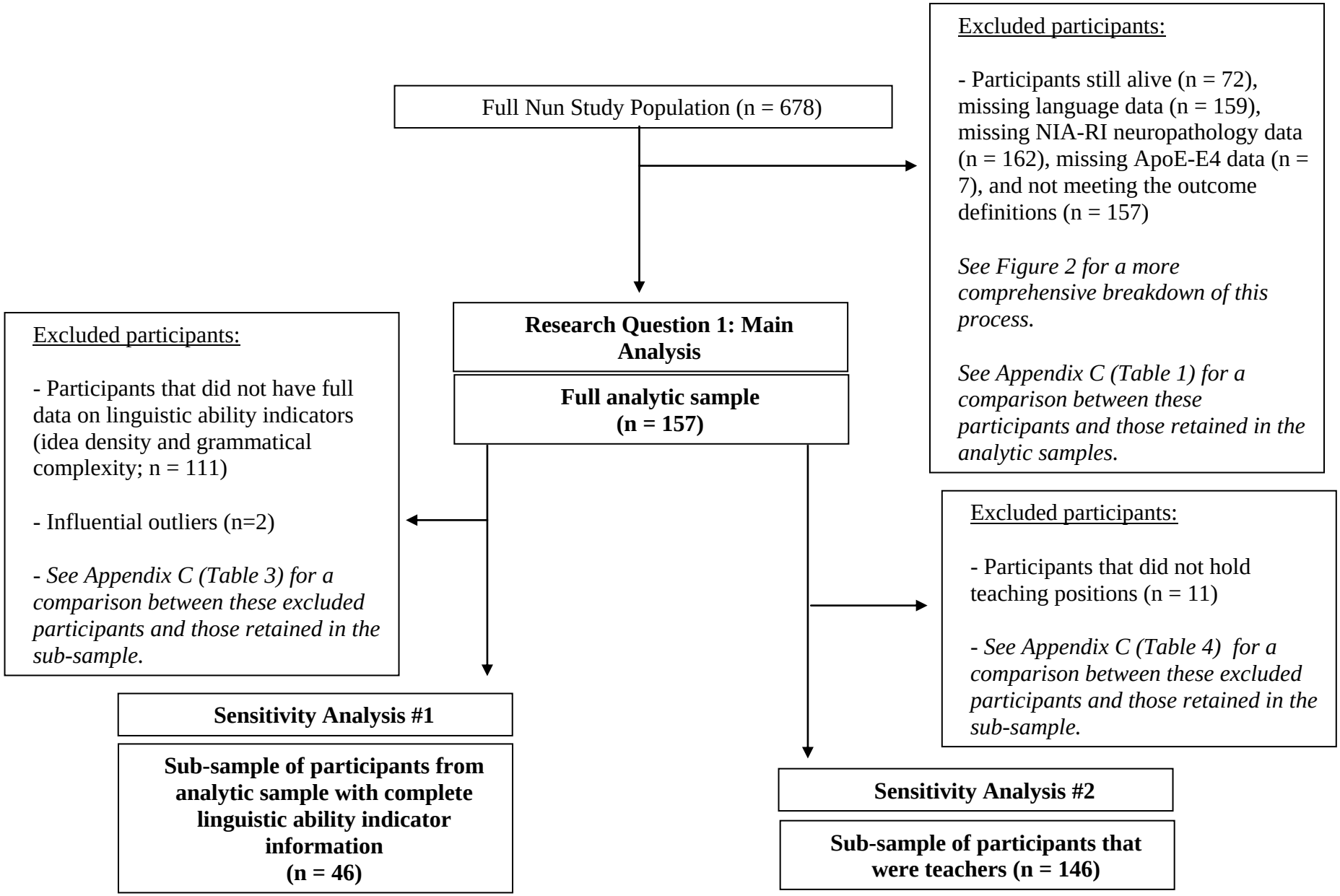

Figure 3. Flow chart outlining the different samples used in Research Question 1. Bolded samples indicate the samples that were employed in logistic regression analyses.
Abbreviations: NIA-RI = National Institute of Aging – Reagan Institute neuropathological criterion; ApoE-E4 = Apolipoprotein E-E4

4.3.2 Research Question 2

4.3.2.1 Analytic Sample Derivation

The main analytic sample used for addressing Research Question 2 was composed of 325 Nun Study participants who had complete cognitive information from baseline and at least one follow-up cognitive assessment. The participants in this sample also were required to have data concerning multilingualism and ApoE-E4 status. Figure 4 outlines how this analytic sample was derived from the original Nun Study population. Participants missing information on one or more variables of interest were excluded from the analytic sample. Participants were also excluded from this analytic sample if they transitioned from clinically demented to non-demented at any point in time during the study (also referred to as displaying "back-transition" behaviour). Some participants displaying this behaviour back-transitioned once before death or study completion while others back-transitioned several times before death or study completion. Since these participants could not be definitively classified as to time of dementia onset and the number of back-transitions varied across the group of participants, all participants displaying this behaviour were excluded from the analyses. Appendix C (Table 5) provides an in-depth comparison of excluded observations to those retained by the Research Question 2 analytic sample. A detailed description of participants displaying "back-transition" behaviour is also included in Appendix C (p. 218).

The participants excluded from this analytic sample were significantly different than the participants retained in the analytic sample across several measures. For instance, participants in the analytic sample were more highly educated than participants excluded from the sample (14.9% of excluded participants had only grade-level education compared to 4.6% of participants in the analytic sample). Excluded participants were also older at baseline cognitive

assessment (mean = 84.2 years; SD = 5.7 years) than participants retained by the analytic sample (mean = 82.4 years; SD = 5.0 years; $p<0.001$). A significantly greater proportion of excluded participants also had at least one ApoE-E4 allele (28.0% versus 18.1% of included participants; $p<0.01$) and held occupations other than teaching positions (15.0% versus 5.5% of included participants; $p<0.001$). Given a greater proportion of excluded participants were more at risk from an ApoE-E4 allele and older ages than participants who were not excluded, it is possible that the present analyses may not have detected all possible associations with outcomes of dementia.

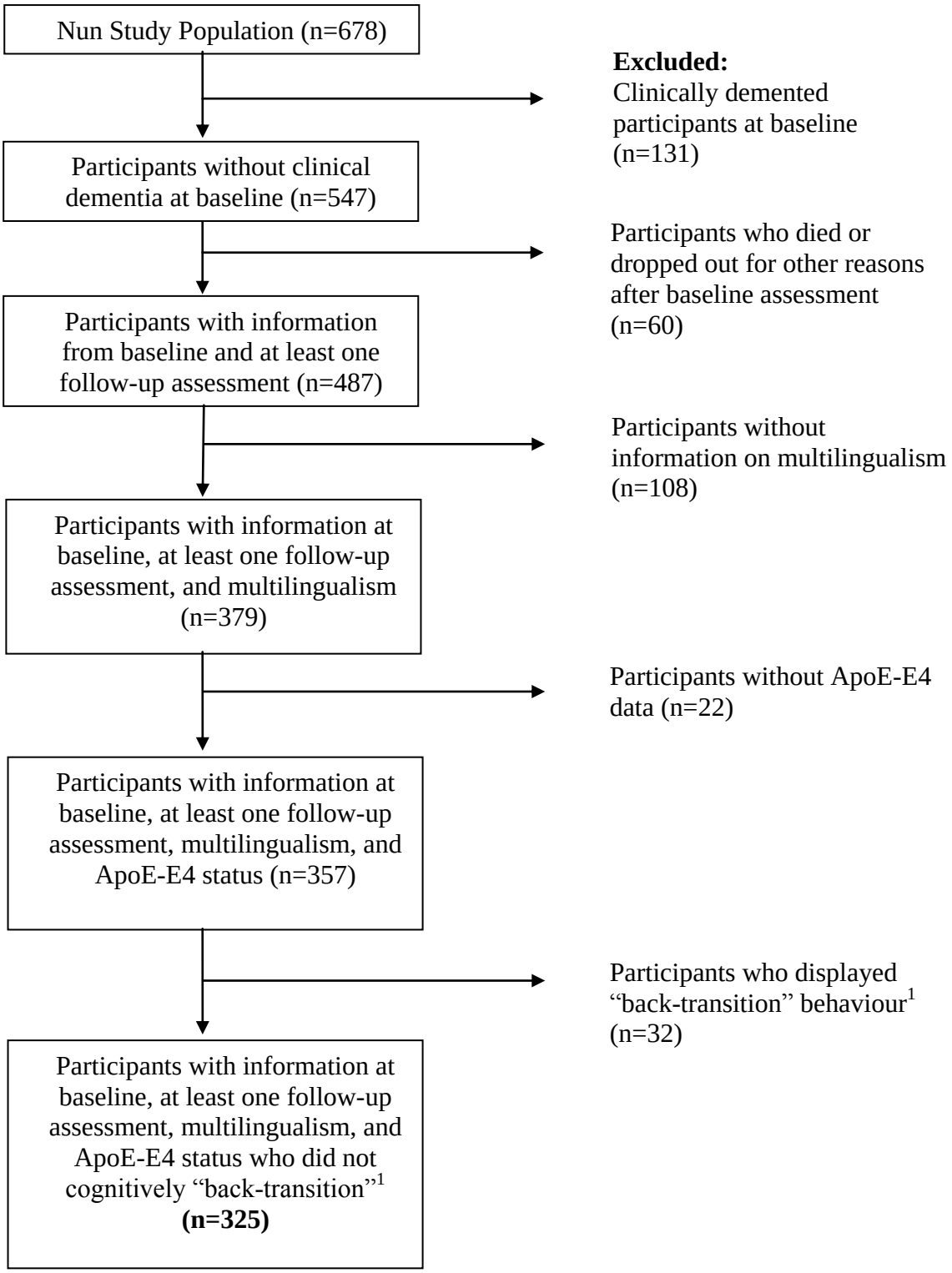

Figure 4. Flow chart outlining the derivation of the analytic sample used for the Research Question 2 main analyses.

[1] Some participants were found to transition back from clinically demented to non-demented throughout the duration of the Nun Study. These participants were consequently referred to as having "back transitions". Since the time of dementia onset of these participants was unclear, they were excluded from the analyses.

Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele

4.3.2.2 Sensitivity Analysis Sub-Sample Derivation

A sensitivity analysis evaluating participant linguistic ability was also conducted to supplement the primary Research Question 2 analyses. This sensitivity analysis included participants from the primary Research Question 2 analytic sample who had complete linguistic ability information; all other participants missing linguistic ability data were excluded. This sub-sample was comprised of 40 participants. A significantly greater proportion of the excluded participants had lower educational levels than the participants included in this sensitivity analysis; 17.7% of excluded participants had educational levels lower than a Bachelor's degree compared with only 3.8% of participants included in the analysis ($p<0.001$). Excluded participants were also older at baseline assessment than included participants (84.4 years versus 79.8 years; $p<0.001$), and a significantly greater proportion of excluded participants held occupations other than teaching positions (12.1% compared to 1.9% of included participants; $p<0.001$). Please refer to Appendix C (Table 6) for additional comparisons between the excluded participants and those included in the analysis.

4.3.3 Variable Selection

The initial selection of covariates for Research Question 1 analyses was guided by findings from the literature review. Only variables available in the Nun Study dataset could be considered for the analyses. All of the same covariates considered for incorporation into the Research Question 1 analyses were also considered for inclusion in the Research Question 2 analyses. The log likelihood test was employed to evaluate which variables were included in the Research Question 2 analyses (see Section 5.2.2 for more details concerning this selection method and analyses).

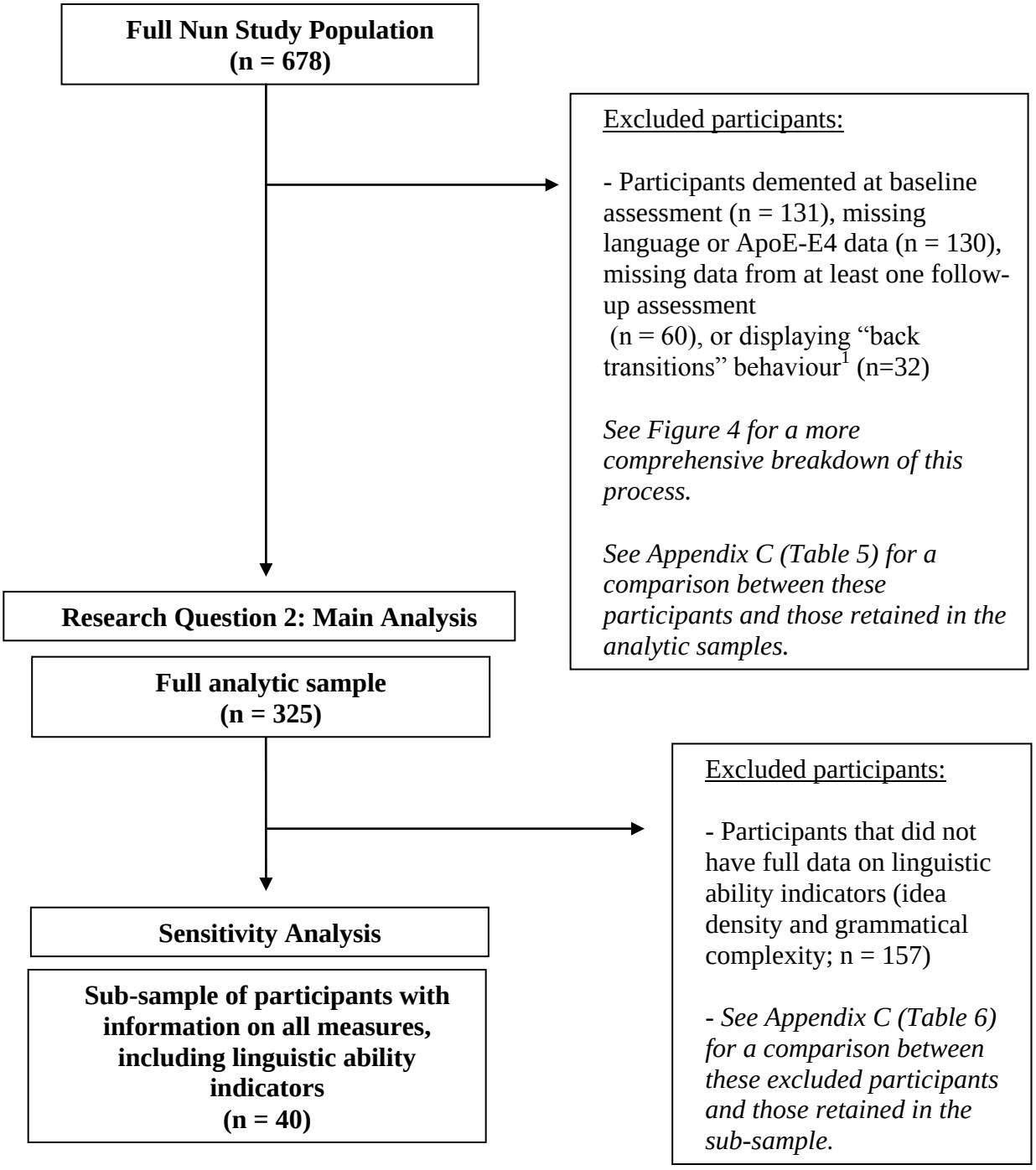

Figure 5. Flow chart outlining the Research Question 2 linguistic ability sensitivity analysis sub-sample derivation.

[1] Some participants were found to transition back from clinically demented to non-demented throughout the duration of the Nun Study. These participants were consequently referred to as having "back transitions". Since the time of dementia onset of these participants was unclear, they were excluded from the analyses.

Abbreviations: ApoE-E4 = Apolipoprotein E-E4

4.4 Measures

4.4.1 Multilingualism

Data on multilingualism were collected by means of self-report questionnaire and were available to the Nun Study through convent archival access. The questionnaire was administered in 1983 (eight years before the Nun Study began), with the intent to facilitate foreign mission placements. Participants were asked to report the languages with which they had proficiency, so that options for international placements could be assessed. The number of languages each participant had proficiency with was ascertained from this report; participants reported proficiencies ranging from one to five languages. All participants spoke English, and there were a variety of additional languages spoken within the sample: 41.1% of participants providing data spoke German, 18.3% spoke French, 12.6% spoke Spanish, 10.8% spoke Polish, 3.2% spoke Italian, 4.9% spoke Latin, while other languages spoken included Czech, Slovak, Japanese, and Chamorro. Aside from this questionnaire, no other formal language proficiency criteria were used to ascertain exposure status in the current study. Please see Appendix D (Table 1) for additional details concerning languages spoken within each analytic sample. It was not recorded which particular languages were spoken as the first and additional languages. The equal use of multiple languages every day (e.g., being "balanced" in many languages) was also not a requirement in order to qualify as being multilingual.

Participants were classified according to the number of languages spoken, with participants fluent in only one language classified as the reference group. The number of categories depended on the analyses employed and the sample sizes in the analyses; since a small proportion of participants spoke four and five languages proficiently, these categories were combined. Therefore, multilingualism was either defined as a four-level variable (with

participants divided by proficiency in one, two, three, and four or more languages), a three-level variable (with participants divided by proficiency in one, two, and three or more languages), or a two-level variable (with participants divided by proficiency in one vs. two or more languages). Any future reference to "participants who spoke x number of languages" should be interpreted as "participants who spoke x number of languages with proficiency".

4.4.2 Alzheimer Disease and Dementia

4.4.2.1 Research Question 1

The outcome of interest for Research Question 1 was neuropathologically confirmed AD (i.e., clinically demented cases that were confirmed post-mortem to have neuropathology characteristic of AD). This outcome was referred to as AD, and any reference to AD in Research Question 1 and its analyses should be interpreted as neuropathologically confirmed unless otherwise specified. AD was diagnosed based on a review of neuropathologic findings, clinical and functional information from annual assessments, and medical records. AD was defined as having a clinical diagnosis of dementia using the DSM-IV criterion (American Psychiatric Association, 1994) and a neuropathological diagnosis of "high likelihood" AD as indicated by the NIA-RI neuropathological criterion (The National Institute on Aging - Reagan Institute (NIA-RI) Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease , 1997). Clinical dementia was diagnosed in participants exhibiting a gradual decline in overall cognitive function, impaired memory (a Delayed Word Recall score <4), impairments in at least one other cognitive domain (Verbal Fluency score <11, Boston Naming <12, or Constructional Praxis <9), and impairments in ADLs. These participants also were required to exhibit functional decline over time.

Participants without AD (the control participants) did not have clinical dementia and either had no neuropathology or neuropathology characteristic of "low likelihood" AD as indicated by the NIA-RI criterion. These participants had cognitive scores within normal limits (within 1.5 standard deviations below the age-appropriate mean) on the previously mentioned cognitive tests, intact global cognitive ability according to the MMSE (Folstein, Folstein, & McHugh, 1975; Kukull et al., 1994), and intact ADLs, as observed from performance-based tests. All cut points used to determine intact cognitive function and impaired function required for dementia were derived from the normative data for the CERAD battery (Riley, Snowdon, & Markesbery, 2002; Welsh, Butters, Hughes, Mohs, & Heyman, 1992).

Details on neuropathological examinations and techniques may be found in previous Nun Study publications (Riley et al., 2002; Snowdon, 1997). Using the NIA-RI criterion, assessments of AD neuropathology were based on counts of NFTs and NPs. This criterion was preferred for this investigation because it considers both AD neuropathology types, whereas other criteria can account for only one type. The NIA-RI neuropathological criterion for "high likelihood" AD was used to define cases of AD, while controls were required to have "low likelihood" AD neuropathology. Participants with dementia but "low" or "intermediate likelihood" neuropathology were excluded from the sample used for addressing Research Question 1, as their dementia may be due to causes other than AD. Those who did not have clinical dementia but had "intermediate" or "high likelihood" neuropathology were also excluded because dementia is a critical component for a diagnosis of AD. By using strict criteria for AD cases and controls, outcome misclassifications were minimized. Please refer to Table 1 for a summary of the outcome definitions.

Table 1. Outcome definitions by Research Question.

Research Question	Outcome Criterion	Case Definition	Control Definition
1	1. Clinical symptomatology	Clinical dementia[a]	Not clinically demented[a]
	2. Neuropathology	NIA-RI (high likelihood[b])	NIA-RI (low likelihood[b])
2	1. Clinical symptomatology	Clinical dementia[a]	Not clinically demented[a]

[a] According to the DSM-IV definition (American Psychiatric Association, 1994). A dementia diagnosis requires impaired cognitive function, memory impairment, impairment in one other cognitive domain, and ADL impairment (functional decline over time must also be demonstrated).
[b] According to the original NIA-RI consensus article (1997).
Abbreviations: AD = Alzheimer disease; NIA-RI = National Institute for Aging, Ronald and Nancy Reagan Institute of Alzheimer's Disease

4.4.2.2 Research Question 2

The outcome used for addressing this Research Question was dementia, which was based on clinical information only. Cases were participants developing dementia, while controls were participants who did not develop dementia. As specified in section 4.4.2.1, clinical diagnoses of dementia were made using the DSM-IV criterion (American Psychiatric Association, 1994) in participants exhibiting a gradual decline in overall cognitive function, impaired memory (a Delayed Word Recall score <4), impairments in at least one other cognitive domain (Verbal Fluency score <11, Boston Naming <12, or Constructional Praxis <9), and impairments in ADLs. A diagnosis of clinical dementia also required the exhibition of functional decline over time. It was not possible to measure the precise time at which AD neuropathology became present in a given participant's brain; therefore, the use of an outcome based on only clinical data was the most practical choice. This analysis estimated the probability of dementia development by considering data from each follow-up cognitive assessment attended by both dementia cases and control participants.

4.4.3 Covariates

The covariates used in this investigation were chosen based on availability in the dataset and their established relationships with AD; these included age, level of formal education, immigration status, ApoE-E4 status, and occupation. Linguistic ability (i.e., idea density and grammatical complexity) covariate data were available in only a sub-set of participants; therefore, sensitivity analyses were performed after the main analyses for each Research Question in order to analyse these covariates. A second sensitivity analysis was performed within the Research Question 1 sample in order to evaluate the influence of the number of years spent working as a teacher. This second Research Question 1 sensitivity analysis utilized a sub-sample of only those participants from the original Research Question 1 analytic sample who held teaching positions.

Age, the most established risk factor for sporadic AD, was measured as a continuous variable. When addressing Research Question 1, age at last cognitive assessment was used, as this best characterized the final cognitive states of both cases and controls. With respect to Research Question 2, age at baseline cognitive assessment was used as the "age" covariate. Study participant ages at baseline assessment ranged from 75-102 years (mean=83.3 years).

Educational level was categorized in the Nun Study as having completed grade school, high school, a Bachelor's degree, or a Master's degree or higher. Since few participants had educational levels lower than a Bachelor's degree, the two lowest educational categories were combined when education was incorporated into regression models. Immigrant status was categorized as a dichotomous variable (yes/no) based on whether a participant was born in the U.S.

ApoE-E4 status was classified as having one or more ApoE-E4 alleles or having none. Occupation was classified on three levels: teacher, house sister, and other. When incorporated into regression models, occupation was categorized dichotomously (teacher or other) as few participants in the sample population occupied positions other than teachers. The total number of years spent as a teacher, when analysed in the occupation sensitivity analysis in Research Question 1, was treated as a continuous variable.

Variables measuring linguistic ability were idea density (Kintsch & Keenan, 1973) and grammatical complexity (Cheung & Kemper, 1992). Idea density was defined as the average number of ideas expressed per ten words. Grammatical complexity scores ranged from zero (simple one-clause sentences) to seven (complex sentences using multiple clauses and embedding). Mean idea density and grammatical complexity scores were calculated for each participant based on the final ten sentences of each autobiography. These scores were then ranked within each convent. The rankings were divided into quartiles for use in these analyses. Only 180 members of the original Nun Study population provided handwritten autobiographies from which these variables were derived; therefore, sensitivity analyses using sub-samples of participants with full linguistic ability data (i.e., data on idea density and grammatical complexity) were performed to supplement the main analyses for each Research Question.

The analysis performed in order to address Research Question 2 also considered each transition period, or the time period between cognitive assessments, as a potential covariate in the analyses. This investigation utilized data from 12 cognitive assessments; therefore, 11 transition periods (mean = 1.51 years in length; SD = 0.32) were evaluated for inclusion in the analyses.

4.5 Ethics

Informed consent from all study participants was originally obtained in 1990 and updated in 2006. Original ethics clearance for the Nun Study was granted by the University of Kentucky; the Nun Study offices have since moved from the University of Kentucky to the University of Minnesota. Data were entered into the database according to participant ID number; a separate, independent set of ID numbers was used for the pathologic data. This investigation used these previously collected data, which were stored at the University of Waterloo in locked cabinets and electronically on password-protected computers. Access was restricted to authorized personnel. Ethics clearance was obtained for this project through the Office of Research Ethics at the University of Waterloo (ORE #16551), and individuals involved in this investigation read and signed confidentiality agreements outlining the ethical protocol of the project (Appendix E).

5.0 Data Analysis

All analyses were performed using SAS version 9.2 (SAS Institute Inc., Cary, North Carolina). Descriptions of the general analytic methods used by this investigation are outlined in the following sections.

5.1 Descriptive Analyses

Exploratory analyses were performed for all variables using univariate and bivariate procedures. The distributions of variables by outcome status for each question are presented in the results section. Additional comparisons between analytic samples and participants excluded from analyses are available in the tables included in Appendix C. In the bivariate analyses, Pearson chi-square tests, with Yates continuity correction and Fisher's exact tests as needed, were used to measure associations between categorical variables. Cochran-Armitage tests for trend were also used to assess associations between the outcome and ordered categorical variables. Independent sample t-tests were used to assess the association between continuous and dichotomous variables. Depending on whether variances of a given variable were deemed to be equal or unequal, either the pooled method (when variances between groups on a given variable were deemed to be equal) or the Satterthwaite method (when variances between groups on a given variable were found to be unequal) was employed.

5.2 Multivariate modelling

5.2.1 Research Question 1

In order to address Research Question 1, the influence of the exposure and covariates on the outcome was assessed using multiple logistic regression procedures. Backward elimination was used as the method of determining which variables were statistically important in the

regression models. This method was preferred over other standard selection methods as it has been shown to yield a lesser mean squared error when compared to forward selection (Kennedy & Bancroft, 1971). The significance (α) levels for variable selection in the backward elimination regression models were 0.15 for main effects and 0.05 for interactions. These significance levels were in accordance with previous recommendations (Lee & Koval, 1997; Tyas, Koval, & Pederson, 2000). When backwards selection procedures were performed for the analyses using the main analytic sample and the sub-sample of teachers, the only variables to be retained by the final regression models were ApoE-E4 status and age at last cognitive assessment. Given the aim of Research Question 1 was to evaluate the association between multilingualism and AD in the context of other covariates, when multilingualism was not retained by the backwards procedure it was forced into the model along with *a priori* variables not meeting the prescribed significance levels (education, immigrant status, and occupation) so that the relationship between multilingualism and AD could be evaluated in a comprehensive manner.

Association strength was assessed using odds ratios (ORs) with 95% confidence intervals (CIs). ORs represent the odds of exposure among cases to the odds of exposure among controls. An OR of one suggests no relationship. An OR of greater than one suggests the exposure is associated with a greater risk of disease than the reference, while an OR of less than one suggests the exposure is associated with a lesser risk of disease (Friis & Sellers, 2009). The profile likelihood-based estimation method for estimating 95% CIs, which are preferred for computing CI estimates with relatively small samples, was used as it allows for asymmetric CI estimates (Evans, Kim, & O'Brien, 1996).

Lack of fit analyses, tests of multicollinearity, and residual diagnostics were performed in order to assess how well the data fit the logistic regression models. Hosmer-Lemeshow goodness-of-fit test statistics were conducted for each model, using the LACKFIT command in PROC LOGISTIC. Models were rejected if the goodness-of-fit statistic p-values were less than 0.05. Tests of multicollinearity among independent variables were executed using the PROC REG procedure in SAS. All models were subjected to an examination of residual diagnostics, which was done using the INFLUENCE and IPLOTS commands in PROC LOGISTIC. The critical value of ±1.96 (corresponding to a 0.05 significance level) was used to determine which observations had significant influences on the fit of their respective models. The same critical value was used when the DFBETA, C and CBAR residual diagnostics were examined. DFBETA values measure the changes in parameter estimates when a given observation was deleted (SAS Institute Inc., 2009), while C and CBAR values indicate how influential observations are on their respective parameter estimates (SAS Institute Inc., 2009).

5.2.2 Research Question 2

Discrete-time survival analysis was used to calculate whether the probability of dementia development varied by multilingualism and time. Discrete-time survival analysis allows for examination of the longitudinal progression of the probability that an event occurs (Muthén & Masyn, 2005). Discrete-time survival analysis does not treat time as a continuous variable, but rather as discrete units or chunks. In this study, cognitive assessments occurred in approximately consistent time intervals. In the Nun Study, the average period of time (i.e., the average transition period length) that elapsed between each cognitive assessment was 1.51 years

(SD = 0.32 years). Please see Appendix F (Table 1) for additional details concerning these discrete time units.

This discrete-time survival analysis calculated the hazard probability of an event (in this case, the development of dementia). The hazard probability relates to the proportion of participants at risk for the event that actually experience the event in a defined time period. A hazard function refers to the chronological pattern of the estimated hazard probabilities over time. The magnitude of a hazard function in a specific time period describes the magnitude of the estimated risk for the event in that time period – the greater the risk, the higher the hazard in the given time period.

The discrete-time survival analysis produced assessments of risk using logistic regression models. Regression models were constructed using the likelihood ratio test, which gauged model suitability by measuring the difference between two models' deviance statistics (deviance is equal to -2 log likelihood). Generally, a model with a lower deviance is thought to better fit the data than a model with a higher deviance. Usually the inclusion of an extra parameter results in a smaller deviance; however, the inclusion of additional parameters reduces a model's statistical power. The likelihood ratio test established whether the inclusion of an additional parameter was justified by an appropriate decrease in deviance. The differences in model deviances were assumed to be distributed, under the null hypothesis, as approximately chi-square (with the appropriate degrees of freedom equal to the difference in the number of parameters in the two nested models being compared). The differences in model deviances were then compared to the appropriate chi-square value at a cut-off point of $p=0.05$; any deviance difference greater than the value at this point indicated that the larger model (with one extra parameter than the model it was compared to) had a significantly better fit with the data.

According to Singer and Willett (2003), this model selection method is preferred for choosing discrete-time survival regression models.

All potential predictors were considered for inclusion in the regression models used for the discrete-time survival analysis. These predictors were all possible categorizations of multilingualism (i.e., classified on four, three, or two levels), all possible time indicators (predictive estimates associated with each of 11 transition periods), occupation, level of educational attainment, age at baseline cognitive assessment, immigrant status, and ApoE-E4 status. Grammatical complexity and idea density were also considered for inclusion in the Research Question 2 sensitivity analysis.

Differences in dementia risk were assessed using hazard function estimates generated from final logistic regression models. ORs were also generated from the final logistic regression models. Final regression models were assessed for goodness-of-fit using the Hosmer-Lemeshow goodness-of-fit test statistic. Models were rejected if the Hosmer-Lemeshow goodness-of-fit test statistic p-values were less than 0.05. Tests of multicollinearity among independent variables were performed and all models were subjected to an examination of residual diagnostics. The critical value of ±1.96 (corresponding to a 0.05 significance level) was used to determine which observations had significant influences on the fit of their respective models.

6.0 Results

6.1 Research Question 1

6.1.1 Full Analytic Sample

6.1.1.1 Descriptive Statistics

In the full analytic sample, 65 participants (Table 2) had AD (41.4%), while the remainder (n=92) was classified as controls (individuals without dementia who did not meet neuropathologic criteria for AD). Overall, 71.3% of the analytic sample spoke more than one language; there were no significant differences in the number of languages spoken between participants according to outcome status. There were no significant differences in educational attainment or immigration status between AD status groups. The participants who developed AD were significantly older than the control participants at the last cognitive assessment (91.2 years versus 89.5 years; $p<0.05$). A significantly greater proportion of participants developing AD were carriers of at least one ApoE-E4 allele, compared with controls (50.8% versus 8.7%; $p<0.001$). The trend between increasing number of ApoE-E4 alleles and increasing AD risk was also found to be significant; no other significant trends were detected.

Table 2. Participant characteristics by AD status (n=157).

Variable		Total (n=157)	Control[1] (n=92)	AD[2] (n=65)
Multilingualism				
# languages (%)	1	28.6	28.2	29.2
	2	50.3	50.0	50.8
	3	17.2	16.3	18.5
	4	1.3	1.1	1.5
	5	2.6	4.4	0.0
2+ languages (%)		71.4	71.8	70.8
Covariates				
Age at last cognitive assessment (years)	mean (SD)	90.2 (5.0)	89.5 (5.0)	91.2 (4.8)*
Education (%)	Grade school	7.0	4.4	10.8
	High school	3.2	2.3	4.6
	Bachelor's degree	40.8	40.2	41.5
	Master's degree +	49.0	53.3	43.1
Occupation (%)	Teacher	93.0	94.6	90.8
	House sister	6.4	4.3	9.2
	Other[3]	0.6	1.1	0.0
Immigrant to USA (%)		4.5	4.4	4.6
1+ ApoE-E4 alleles (%)		26.1	8.7	50.8***

*p<0.05; **p<0.01; ***p<0.001

[1] Defined as not having clinical dementia and having no neuropathology/ "low likelihood" neuropathology according to the NIA-RI criterion (The National Institute on Aging- Reagan Institute (NIA-RI) Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease , 1997).

[2] Defined as exhibiting symptoms consistent with clinical dementia and having neuropathology classified as "high likelihood" by the NIA-RI criterion (The National Institute on Aging - Reagan Institute (NIA-RI) Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease , 1997).

[3] An example of another occupation held by a given participant was a nurse's aide.

Abbreviations: AD = Alzheimer disease; ApoE-E4 = Apolipoprotein E-E4 allele; NIA-RI = National Institute of Aging – Reagan Institute; SD = standard deviation

6.1.1.2. Multivariate Logistic Regression Models

Logistic regression models were initially developed using the selection methods outlined in Section 5.2.1. None of the interactions between covariates and multilingualism were found to be significant. ApoE-E4 allele status and age at last cognitive assessment were the only covariates deemed significant for model retention. Since the aim of this investigation was to evaluate the relationship between AD and multilingualism in the context of other covariates unique to this study population, alternative regression models including all variables of interest were analyzed rather than only models generated by backward elimination. All regression models can be found in Appendix G (Table 1).

Multilingualism (defined as a four-level variable: speaking one, two, three, or four or more languages) was not significantly associated with AD in either crude or adjusted logistic regression models (Table 3). The ORs produced at each level of the exposure in the crude model were not statistically significant but displayed an interesting trend. Participants who spoke one language had odds of developing AD similar to participants speaking two (OR = 0.98; 95% CI = 0.47- 2.07) or three languages (OR = 1.09; 95% CI = 0.41-2.87). Conversely, participants speaking four or more languages had lower odds of developing AD compared to monolinguals (OR = 0.27; 95% CI = 0.01-1.88). A similar trend was evident across exposure levels in the adjusted model. The possession of at least one ApoE-E4 allele was the only significant AD risk factor in the adjusted model (OR=12.26; 95% CI = 4.88-30.80), while increasing age at last cognitive assessment was also associated with a marginally significant increase in AD odds (OR=1.07; 95% CI = 0.99-1.17). Higher educational attainment, occupying a teaching position, and immigrant status all appeared to reduce the likelihood of AD according to this model, but these relationships did not reach statistical significance.

Similar results were found when multilingualism was instead defined on two levels (two or more languages versus one; Table 4). According to the unadjusted model, speaking two or more languages was not significantly associated with the likelihood of AD compared to speaking only one language (OR=0.95; 95% CI = 0.47-1.92). A similar estimate of AD odds was generated from the adjusted model for participants speaking two or more languages compared to those speaking only one (OR = 1.05; 95% CI = 0.45-2.50). Similar to the results from the previous adjusted regression model, the only variable significantly associated with increased AD odds was the possession of an ApoE-E4 allele (OR=12.41; 95% CI = 5.21-32.86).

Table 3. Association between Alzheimer disease and multilingualism using a four-level multilingualism variable.

Model	Exposure	OR	95% CI
Crude	Multilingualism		
	Speaking two languages vs. one	0.98	0.47, 2.07
	Speaking three languages vs. one	1.09	0.41, 2.87
	Speaking four or more languages vs. one	0.27	0.01, 1.88
Adjusted	Multilingualism		
	Speaking two languages vs. one	0.99	0.40, 2.46
	Speaking three languages vs. one	1.39	0.43, 4.50
	Speaking four or more[1] languages vs. one	0.61	0.06, 6.15
	Age at last cognitive assessment (per year increase)	1.07	0.99, 1.17
	Level of attained formal education		
	Bachelor's degree vs. high school or less	0.60	0.08, 4.38
	Master's degree or higher vs. high school or less	0.43	0.06, 3.18
	Occupation		
	Teacher vs. other[2]	0.83	0.09, 7.42
	Immigrant status		
	Immigrant vs. US-born	0.72	0.10, 5.00
	ApoE-E4 status		
	Possessing an ApoE-E4 allele	**12.26**	**4.88, 30.80**

Bolded values indicate statistically significant results.
[1] Maximum number of languages spoken was five; participants speaking five languages (n=4) were grouped together with those speaking four due to limited numbers.
[2] Examples of other occupations included house sisters and nurse's aides.
Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele; OR = odds ratio estimate; 95% CI = 95% confidence interval

Table 4. Association between Alzheimer disease and multilingualism using a two-level multilingualism variable.

Model	Exposure	OR	95% CI
Crude	*Multilingualism*		
	Speaking two or more languages vs. one	0.95	0.47, 1.92
Adjusted	*Multilingualism*		
	Speaking two or more languages vs. one	1.05	0.45, 2.50
	Age at last cognitive assessment (per year increase)	1.07	0.99, 1.17
	Level of attained formal education		
	Bachelor's degree vs. high school or less	0.59	0.07, 4.02
	Master's degree or higher vs. high school or less	0.43	0.05, 2.90
	Occupation		
	Teacher vs. other[1]	0.87	0.11, 9.40
	Immigrant status		
	Immigrant vs. US-born	0.70	0.09, 4.56
	ApoE-E4 status		
	Possessing an ApoE-E4 allele	**12.41**	**5.21, 32.86**

Bolded values indicate statistically significant results.
[1] Examples of other occupations included house sisters and nurse's aides.
Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele; OR = odds ratio estimate; 95% CI = 95% confidence interval

6.1.2. Sensitivity Analysis using Linguistic Ability Sub-Sample

While it was ensured that all participants included in the analytic sample had data for every primary covariate of interest, a sub-set of supplementary covariates of interest with analytic value were also considered (i.e., linguistic ability variables and duration of teaching career). Since not all participants included in the original analytic sample had complete data for these additional variables, two sensitivity analyses were performed in order to supplement the primary analyses involved in addressing Research Question 1 (please refer to Section 4.3.1.2 for details concerning the derivation of these analytic samples). The first sensitivity analysis was performed in order to evaluate the relationship between AD and multilingualism in the context of linguistic ability (see Sections 6.1.2.1 and 6.1.2.2), while the second sensitivity analysis was performed in order to evaluate the relationship between AD and multilingualism in the context of teaching career length (see Section 6.1.3).

6.1.2.1 Descriptive Statistics

Table 5 presents the characteristics of the linguistic ability sub-sample by AD status. There were 12 participants with AD (12/46; 26.1%), while the remaining participants (n=34) were controls. Overall, 71.7% (33/46) of the entire sub-sample spoke more than one language: 76.5% (26/34) of control participants spoke more than one language compared to 58.3% (7/12) of participants with AD. No participants with AD spoke more than three languages. While the language differences between AD status categories were apparent, they were not statistically significant. Differences between AD status categories with respect to ApoE-E4 status were, however, statistically significant as 11.8% (4/34) of control participants had an ApoE-E4 allele compared to 75.0% (9/12) of participants with AD. The trend of increasing number of ApoE-E4 alleles and AD risk was also significant. All participants in this sub-sample were teachers that

had been born in the USA (non-immigrants). None of the participants in this sub-sample had educational levels lower than high school completion. With respect to linguistic ability, a significantly greater proportion of participants with AD (58%; 7/12) were classified in the lowest quartile of idea density than control participants (14.7%; 5/34). A greater proportion of control participants had idea densities ranking in the highest quartiles than participants with AD (58.8% vs. 41.7% ranked in the highest two quartiles of idea density). Differences in grammatical complexity between participants in AD status groups followed a pattern similar to that found with idea density; half (6/12) of participants with AD, compared to 17.6% (6/34) of control participants, had grammatical complexity rankings in the lowest quartile.

Table 5. Participant characteristics by AD status: linguistic ability sensitivity analysis (n=46).

Variable		Total (n=46)	Control[1] (n=34)	AD[2] (n=12)
Multilingualism				
# of languages (%)	1	28.3	23.6	41.7
	2	45.6	50.0	33.3
	3	23.9	23.5	25.0
	4	0.0	0.0	0.0
	5	2.2	2.9	0.0
2+ languages (%)		71.7	76.4	58.3
Covariates				
Age at last cognitive assessment (years)	mean (SD)	87.7 (3.7)	87.8 (4.1)	87.2 (2.3)
Education (%)	Grade school	0.0	0.0	0.0
	High school	4.3	0.0	16.7
	Bachelor's degree	43.5	47.1	33.3
	Master's degree +	52.2	52.9	50.0
Immigrant to USA (%)		0.0	0.0	0.0
1+ ApoE-E4 alleles (%)		28.3	11.8	75.0***
Occupation (%)	Teacher	100.0	100.0	100.0
Idea density quartile (%)	1 (low)	26.1	14.7	58.3**
	2	19.6	26.5	0.0
	3	26.1	29.4	16.7
	4 (high)	28.3	29.4	25.0
Grammatical complexity Quartile (%)	1 (low)	26.1	17.7	50.0*
	2	32.6	35.3	25.0
	3	17.4	20.6	8.3
	4 (high)	23.9	26.5	16.7

*p <0 .05, **p<0.01, ***p<0.001

[1] Defined as not having clinical dementia and having no neuropathology/ "low likelihood" neuropathology (The National Institute on Aging - Reagan Institute (NIA-RI) Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease , 1997)

[2] Defined as having clinical dementia and "high likelihood" Alzheimer neuropathology (The National Institute on Aging - Reagan Institute (NIA-RI) Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease Alzheimer's Association Workgroup, 1997)

Abbreviations: AD = Alzheimer disease; ApoE-E4 = Apolipoprotein E-E4 allele; NIA-RI = National Institute of Aging - Reagan Institute; SD = standard deviation

6.1.2.2 Multivariate Logistic Regression Models

Logistic regression models were developed using the selection methods outlined in Section 5.2.1. None of the interactions between covariates and multilingualism were found to be significant. The main effects that met the required significance level for model retention differed depending on the exposure definition used in the model. Akin to the previous regression models, two different categorical definitions of multilingualism were used in these analyses: a definition that classified number of languages spoken into three levels (one, two, and three languages or more), and a two-level definition (one vs. two or more languages). A four-level exposure definition was not employed as the number of participants in this sub-sample was limited. When multilingualism was classified according to the three-level definition, the variables deemed to be significant for model retention were multilingualism, age at last cognitive assessment, and ApoE-E4 status. A logistic regression model including these variables was subsequently analysed (all regression models can be found in tables included in Appendix H), but it was apparent that, when the three-level definition of multilingualism was used, the sub-sample had insufficient observations to reliably produce parameter estimates and 95% CIs. Therefore, ORs were not calculated using this model and multilingualism definition. Instead, a descriptive contingency table (Table 6) and bar graphs (Figures 6 and 7) were constructed in order to illustrate the observed associations between significant covariates, multilingualism, and AD.

Table 6. Contingency table displaying results of the linguistic ability sensitivity analysis using a three-level multilingualism variable.

	Outcome (n; %)	
Exposure	**Control (n=34)**	**AD (n=12)**
Multilingualism		
Speaking one language	8 (23.5%)	5 (41.7%)
Speaking two languages	17 (50.0%)	4 (33.3%)
Speaking three or more[1] languages	9 (26.5%)	3 (25.0%)
Age at last cognitive assessment (years) (mean, SD)	87.8 (4.1)	87.2 (2.3)*
ApoE-E4 status		
No ApoE-E4 allele	30 (88.2%)	3 (25.0%)***
Possessing an ApoE-E4 allele	4 (11.8%)	9 (75.0%)***

* $p<0.05$; **$p<0.01$; ***$p<0.001$

[1]None of the participants in this sub-sample spoke four languages; 1 control participant spoke five languages.

Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele; AD = Alzheimer disease

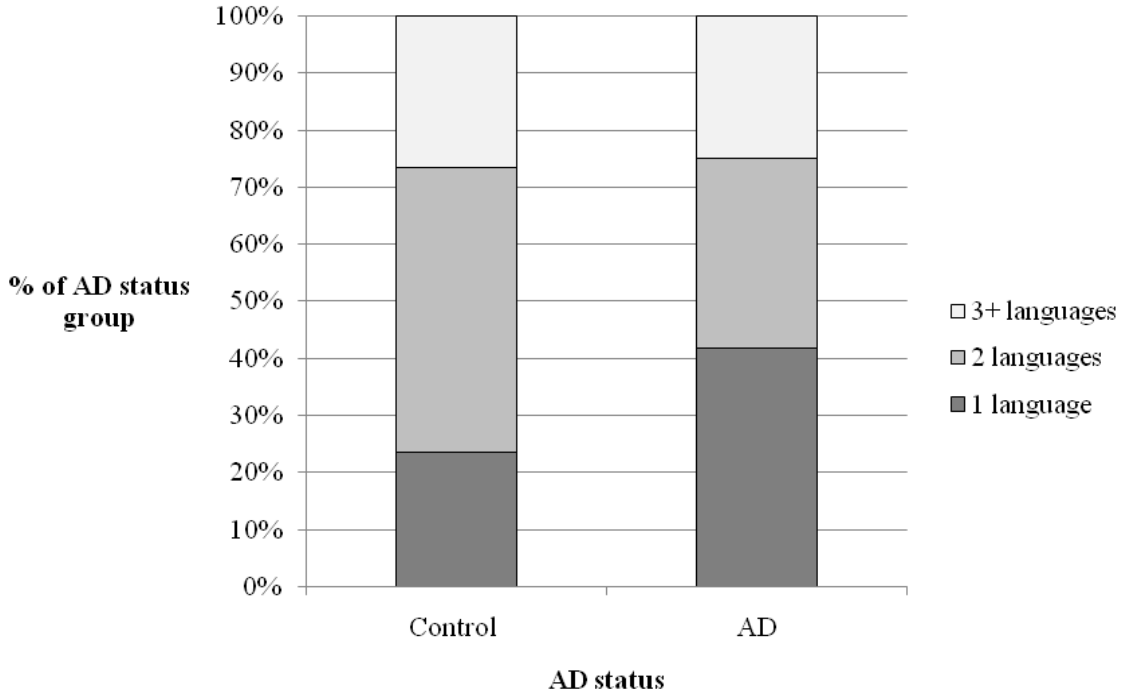

Figure 6. Proportion of each outcome group speaking 1, 2, or 3+ languages: linguistic ability sensitivity analysis.
Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele; AD = Alzheimer disease

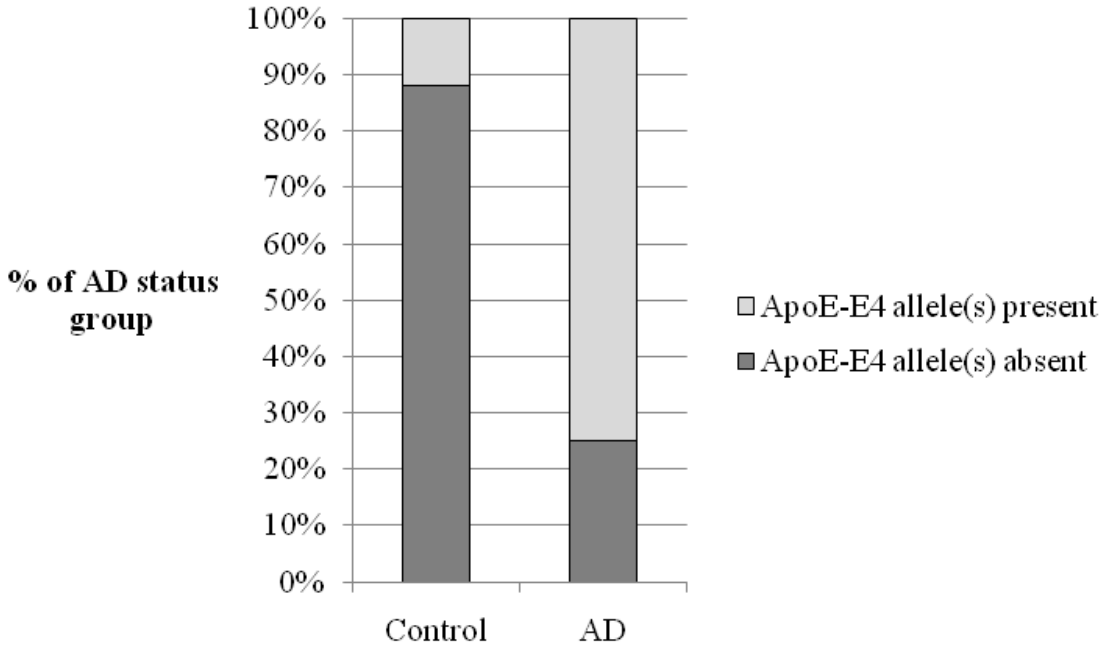

Figure 7. Proportion of each outcome group possessing an ApoE-E4 allele: linguistic ability sensitivity analysis.
Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele; AD = Alzheimer disease

The results displayed in Table 6 and Figure 6 illustrate that a higher proportion of participants who developed AD in this sub-sample were monolingual, while a greater proportion of control participants were proficient in multiple languages. These differences in proportions, however, were not statistically significant. Age at last cognitive assessment and ApoE-E4 status were both significantly associated with AD, which was consistent with the results from the full analytic sample. Using this exposure definition and regression model, ApoE-E4 status yielded the most parameter estimate and separation problems; therefore, the distribution of participant ApoE-E4 genotypes within this sub-sample was examined further (Table 7 and Figure 8). A clear pattern of AD risk was evident from this ApoE-E4 genotype distribution: ApoE-E2 alleles appeared to confer protection against AD, while AD risk increased according to the number of ApoE-E4 alleles. This genotypic distribution also illustrated a reason for the observed separation problems: no individuals with an ApoE-E2 allele developed AD, while no individuals homozygous for the E4 allele remained AD-free.

Table 7. Distribution of apolipoproteinE genotype by Alzheimer disease status: linguistic ability sensitivity analysis.

Exposure	**Outcome (n; %)**	
	Control (n=34)	**AD (n=12)**
ApoE-E4 genotype		
22	1 (2.9%)	0 (0.0%)
23	7 (20.6%)	0 (0.0%)
33	22 (64.7%)	3 (25.0%)*
34	4 (11.8%)	8 (66.6%)***
44	0 (0.0%)	1 (8.3%)

* p<0.05; **p<0.01; ***p<0.001
Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele; AD = Alzheimer disease

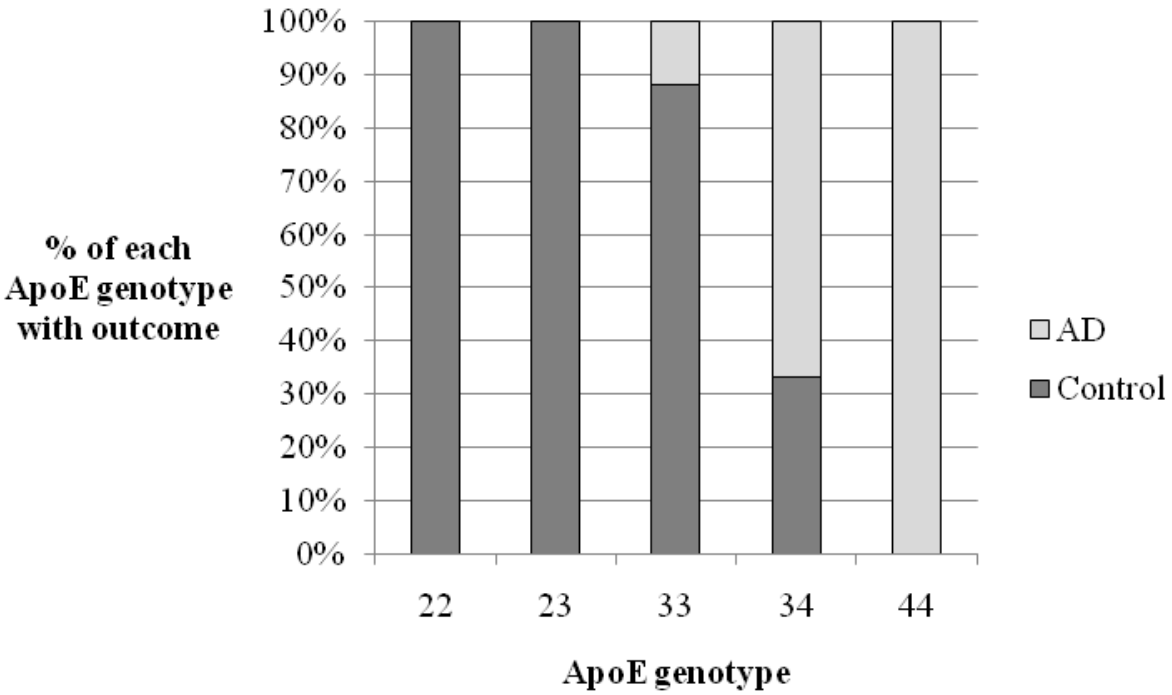

Figure 8. Distribution of ApoE genotype by AD status: linguistic ability sensitivity analysis.
Abbreviations: ApoE = Apolipoprotein E; AD = Alzheimer disease

While the interaction between multilingualism and ApoE-E4 status was not statistically significant because of the limited power of this test, a table was constructed in order to examine the association between multilingualism and AD after stratification by ApoE-E4 status (Table 8). This table illustrated a significant association between multilingualism and AD, but only in participants without an ApoE-E4 allele. Multilingualism was not significantly associated with decreased AD risk in participants possessing one or more ApoE-E4 alleles.

Table 8. Multilingualism by Alzheimer disease status, stratified by apolipoproteinE-E4: linguistic ability sensitivity analysis.

	Outcome (n; %)	
No ApoE-E4 allele	**Control (n=30)**	**AD (n=3)**
Multilingualism*		
Speaking one language	7 (23.3%)	3 (100.0%)
Speaking two languages	14 (46.6%)	0 (0.0%)
Speaking three or more languages	9 (30.0%)	0 (0.0%)
	Outcome (n; %)	
At least one ApoE-E4 allele	**Control (n=4)**	**AD (n=9)**
Multilingualism		
Speaking one language	1 (25.0%)	2 (22.2%)
Speaking two languages	3 (75.0%)	4 (44.4%)
Speaking three or more languages	0 (0.0%)	3 (33.3%)

*$p<0.05$, as determined using Fisher's exact test.
Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele; AD = Alzheimer disease

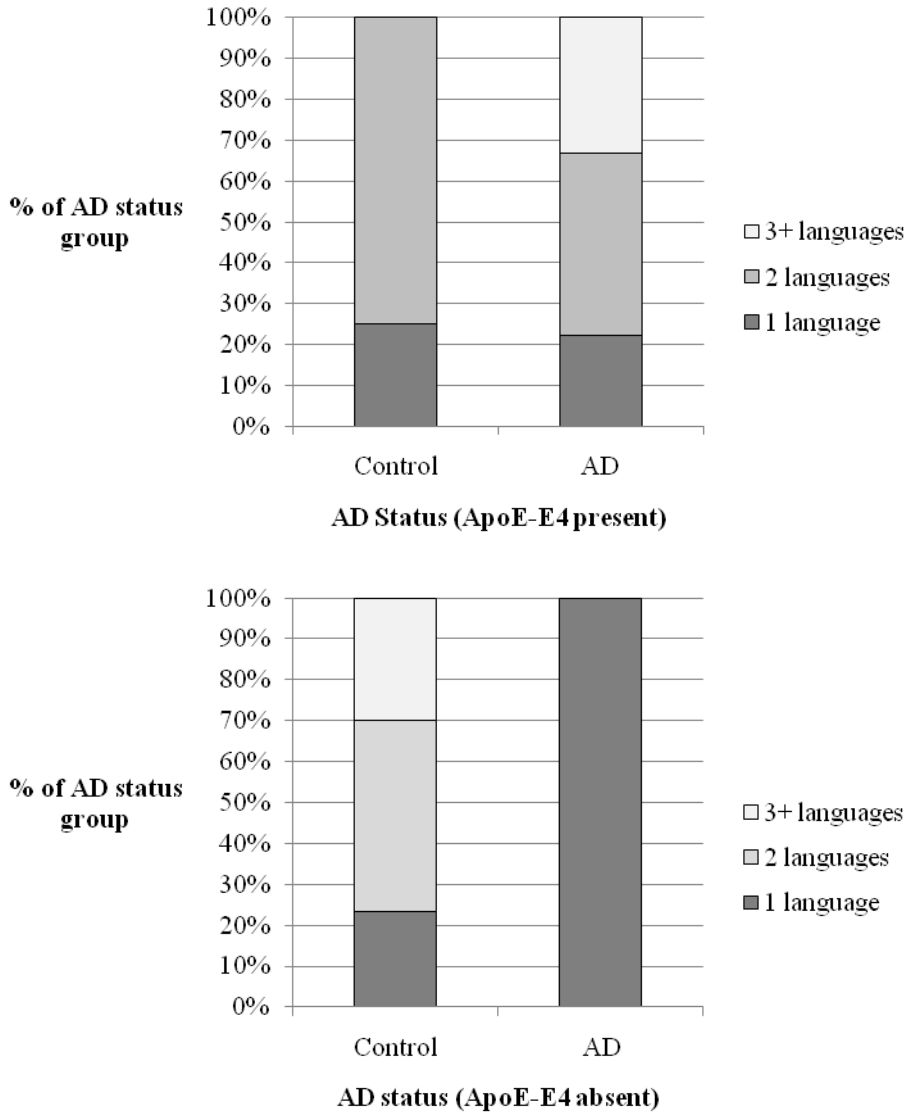

Figure 9. Proportion of each outcome group speaking multiple languages, stratified by ApoE-E4 status: linguistic ability sensitivity analysis.

Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele; AD = Alzheimer disease

When multilingualism was classified according to a two-level definition rather than a three-level definition, multilingualism and grammatical complexity were both found to be significant. A logistic regression model incorporating these variables was subsequently analysed.

Multilingualism was associated with a reduced odds of AD (Table 9), although this association did not meet statistical significance (OR = 0.25; 95% CI = 0.04-1.23). Increased AD risk was not associated with descending quartiles of grammatical complexity. Using the first quartile (quartile of lowest grammatical complexity scores) as the reference group, participants in the third quartile had the largest and only statistically significant reduction in AD likelihood (OR = 0.09; 95% CI = 0.00-0.85). Participants in the second (OR = 0.19; 95% CI = 0.03-1.08) and fourth (OR = 0.14; 95% C.O. = 0.01-1.00) quartiles of grammatical complexity exhibited associations with smaller and non-significant reductions in AD odds.

Table 9. Association between Alzheimer disease and a two-level multilingualism variable, controlling for grammatical complexity: linguistic ability sensitivity analysis.

Exposure	**OR**	**95% CI**
Multilingualism		
Speaking two or more languages vs. one	0.25	0.04, 1.23
Grammatical complexity (quartile[1])		
2 vs. 1	0.19	0.03, 1.08
3 vs. 1	**0.09**	**0.00, 0.85**
4 vs. 1	0.14	0.01, 1.00

Bolded values indicate statistically significant results.

[1] Where one represents the lowest quartile with respect to grammatical complexity.

Abbreviations: OR = odds ratio estimate; 95% CI = 95% confidence interval

Given that the results displayed in Table 8 suggested that an association between multilingualism and AD may differ according to ApoE-E4 status, contingency tables stratified by ApoE-E4 status were constructed to analyze the association between AD and a two-level multilingualism definition. The pattern of AD risk apparent from these tables (not shown) was similar to that seen when multilingualism was defined on three levels: multilingualism appeared to be more strongly associated with lower AD risk in participants without an ApoE-E4 allele than in those with an ApoE-E4 allele. Logistic regression models using a two-level multilingualism definition and stratified by ApoE-E4 status were also analyzed; however, these models could not produce reliable odds ratio estimates, given the small sample size.

6.1.3 Sensitivity Analysis in Teacher Sub-Sample

6.1.3.1 Descriptive Statistics

Table 10 presents the characteristics of the sub-sample restricted to teachers, by AD status. There were 59 participants with AD in this sample (40.4%). Overall, 70.5% (n=103) of the analytic sample spoke more than one language. There were no significant language differences between participants according to AD status. The average teaching career duration was 42.7 years (SD = 8.0 years), and duration of teaching career did not significantly differ between AD cases (mean = 42.9 years; SD = 8.7 years) and controls (mean = 42.6 years; SD = 7.6 years). ApoE-E4 status was the only variable significantly associated with AD, and a significant trend was again detected between ApoE-E4 status and AD.

Table 10. Participant characteristics by AD status: sensitivity analysis restricted to teachers (n = 146).

Variable		Total (n=146)	Control[1] (n=87)	AD[2] (n=59)
Multilingualism				
# of languages (%)	1	29.5	27.7	32.2
	2	47.9	49.4	45.8
	3	18.5	17.2	20.3
	4	1.4	1.1	1.7
	5	2.7	4.6	0.0
2+ languages (%)		70.5	72.3	67.8
Covariates				
Age at last cognitive assessment (years)	mean (SD)	89.9 (4.9)	89.4 (4.9)	90.8 (4.7)
Education (%)	Grade school	0.7	0.0	1.7
	High school	3.4	2.3	5.1
	Bachelor's degree	43.8	42.5	45.8
	Master's degree +	52.0	55.2	47.5
Immigrant to USA (%)		3.4	3.4	3.4
1+ ApoE-E4 alleles (%)		26.7	9.2	52.5***
Total years as a teacher	mean (SD)	42.7 (8.0)	42.6 (7.6)	42.9 (8.7)

*p <0 .05, **p<0.01, ***p<0.001

[1] Defined as not having clinical dementia and having no neuropathology/ "low likelihood" neuropathology according to the NIA-RI criterion (The National Institute on Aging - Reagan Institute (NIA-RI) Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease, 1997).

[2] Defined as exhibiting symptoms consistent with clinical dementia and having neuropathology classified as "high likelihood" by the NIA-RI criterion (The National Institute on Aging - Reagan Institute (NIA-RI) Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease, 1997).

Abbreviations: AD = Alzheimer disease; ApoE-E4 = Apolipoprotein E-E4 allele; NIA-RI = National Institute of Aging – Reagan Institute; SD = standard deviation

6.1.3.2 Multivariate Logistic Regression Models

None of the interaction terms were judged to be statistically significant. The only significant main effects were ApoE-E4 allele status and age at last cognitive assessment. In addition, regression models incorporating all potential covariates were analyzed, so that the relationship between multilingualism and AD could be more comprehensively described in the context of covariates unique to the Nun Study.

Irrespective of whether a four-level or a two-level multilingualism definition was used in the regression analyses, multilingualism was not significantly associated with AD development. When multilingualism was defined on four levels, participants speaking four or more languages were the only multilinguals to demonstrate reduced odds for AD development (OR = 0.60; 95% CI = 0.03-4.66); however, this measure of association was not statistically significant. When a two-level definition was employed, multilingualism was not associated with any significant change in AD odds, compared to monolingual participants (OR = 1.05; 95% CI = 0.45-2.50). Teaching career length also was not significantly associated with the odds of AD development; this was true regardless of whether the four-level multilingualism definition (OR=1.00; 95% CI = 0.96-1.05) or the two-level multilingualism definition (OR = 1.00; 95% CI = 0.96-1.04) was used. Most of the other results from these logistic regression models were similar to those of the full analytic sample, which included participants of all occupations. ApoE-E4 status was the only covariate significantly related to AD odds. Possessing at least one ApoE-E4 allele increased the odds of AD by approximately 12-fold, according to both the model defining multilingualism on four levels (OR = 12.15; 95% CI = 5.08-32.39) and two levels (OR = 12.33; 95% CI = 5.20-32.56).

Table 11. Association between Alzheimer disease and a four-level multilingualism variable: sensitivity analysis restricted to teachers.

Exposure	**OR**	**95% CI**
Multilingualism		
Speaking two languages vs. one	1.00	0.40, 2.46
Speaking three languages vs. one	1.38	0.43, 4.49
Speaking four or more[1] languages vs. one	0.60	0.03, 4.66
Age at last cognitive assessment (per year increase)	1.07	0.99, 1.17
Level of attained formal education[2]		
Bachelor's degree	0.49	0.06, 3.35
Master's degree +	0.36	0.04, 2.38
Teaching career length (per additional year)	1.00	0.96, 1.05
Immigrant status		
Immigrant vs. US-born	0.74	0.09, 4.91
ApoE-E4 status		
Possessing an ApoE-E4 allele	**12.15**	**5.08, 32.39**

Bolded values indicate statistically significant results.
[1] Maximum number of languages spoken was five; participants speaking five languages (n=4) were grouped together with those speaking four due to limited numbers.
[2] Where the reference group consisted of participants with high school education or less.
Abbreviations: AD = Alzheimer disease; ApoE-E4 = Apolipoprotein E-E4 allele; OR = odds ratio estimate; 95% CI = 95% confidence interval

Table 12. Association between Alzheimer disease and a two-level multilingualism variable: sensitivity analysis restricted to teachers.

Exposure	**OR**	**95% CI**
Multilingualism		
Speaking two or more languages vs. one	1.05	0.45, 2.50
Age at last cognitive assessment (per year increase)	1.07	0.99, 1.17
Level of attained formal education[1]		
Bachelor's degree	0.53	0.06, 3.45
Master's degree +	0.39	0.05, 2.34
Teaching career length (per additional year)	1.00	0.96, 1.04
Immigrant status		
Immigrant vs. US-born	0.71	0.09, 4.64
ApoE-E4 status		
Possessing an ApoE-E4 allele	**12.33**	**5.20, 32.56**

Bolded values indicate statistically significant results.
[1]Where the reference group consisted of participants with high school education or less.
Abbreviations: AD = Alzheimer disease; ApoE-E4 = Apolipoprotein E-E4 allele; OR = odds ratio estimate; 95% CI = 95% confidence interval

6.2 Research Question 2

6.2.1 Full Analytic Sample

6.2.1.1 Descriptive Statistics

The sample used for the primary Research Question 2 discrete-time survival analyses contained 325 participants; of these, 33.5% (n=109) developed clinical dementia at some point during 11 waves of follow-up (Table 13). Participants with and without dementia were generally similar in terms of number of languages spoken. Few participants who developed dementia spoke more than three languages (0.9%; n=1) compared to 7.4% (n=16) of participants who remained dementia-free; however, this was not a significant finding (Table 13). Participants were similar with respect to educational attainment: overall, only 8.9% of the entire sample had educational levels lower than a Bachelor's degree. Participants with and without dementia were also similar with respect to occupation (94.0% of controls and 95.4% of participants with dementia were teachers) and immigration status (6.0% of control and 4.6% of participants with dementia were immigrants). Participants who went on to develop dementia were significantly older at baseline assessment (mean = 83.8 years; SD = 5.4 years) than participants remaining dementia-free (mean = 81.7 years; SD = 4.7 years). A significantly greater proportion of participants with dementia had an ApoE-E4 allele compared to participants without dementia (25.7% versus 14.3%, respectively).

Table 13. Participant characteristics by dementia status (n = 325).

Variable		Total (n=325)	Control[1] (n=216)	Clinically demented[2] (n=109)
Multilingualism				
# of languages (%)	1	26.8	27.8	24.8
	2	52.6	48.6	60.5
	3	15.4	16.2	13.8
	4	3.1	4.2	0.9
	5	2.1	3.2	0.0
2+ languages (%)		73.2	72.2	75.2
Covariates				
Age at baseline assessment (years)	mean (SD)	82.4 (5.0)	81.7 (4.7)	83.8 (5.4)***
# of follow-up assessments	mean (SD)	5.3 (3.4)	5.9 (3.5)	4.1 (2.8)***
Education (%)	Grade school	4.6	4.6	4.6
	High school	4.3	3.2	6.4
	Bachelor's degree	37.8	36.6	40.4
	Master's degree +	53.2	55.6	48.6
Occupation (%)	Teacher	94.5	94.0	95.4
	House sister	3.7	4.2	2.7
	Other[3]	1.8	1.8	1.8
Immigrant to USA (%)		5.5	6.0	4.6
Possessing an ApoE-E4 allele (%)		18.1	14.3	25.7*

*p<0.05; **p<0.01; ***p<0.001

[1] Participants in this category did not have clinical dementia (according to the DSM-IV criterion) at any cognitive assessment during the Nun Study.

[2] Participants in this category were dementia-free at baseline cognitive assessment but ultimately demonstrated cognitive states consistent with clinical dementia (according to the DSM-IV criterion).

[3] An example of another occupation held by participants was a nurse's aide.

Abbreviations: ApoE-E4 = Apolipoprotein E-E4, SD = standard deviation

The proportion of participants within the analytic sample developing dementia was similar across each transition period (Table 14). Generally, a greater proportion of participants developing dementia in a given transition period had an ApoE-E4 allele, compared to participants who did not develop dementia in the same transition period (Table 15). Participants who developed dementia in a given transition period were also older than participants not developing dementia in the same transition period. When participants in each dementia status group were stratified by number of languages spoken (Table 16), a smaller proportion of participants speaking four or more languages developed dementia in each transition period.

Table 14. Percentage[1] of sample developing dementia in each transition period.

Outcome (%)	**Transition period**										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Dementia	6.2	4.5	6.9	5.8	6.7	5.1	3.2	4.9	4.6	5.3	5.3
No dementia	85.8	87.7	78.3	85.2	78.8	76.6	80.5	82.2	84.9	77.3	77.2

[1] When columns do not total 100 percent, the remainder of participants died during the given transition period.

Table 15. Participant characteristics, by dementia status, across each transition period.

Participants with dementia	**Transition period**										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
% with E4 allele	22.7	21.4	36.8	38.5	30.8	12.5	25.0	20.0	0.0	25.0	0.0
Current age (mean)	87.4	85.9	88.5	90.9	90.3	93.3	91.3	92.2	95.4	96.0	97.1
Participants without dementia	**Transition period**										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
% with E4 allele	17.8	18.0	14.3	12.1	11.2	10.7	10.1	9.6	11.0	8.6	6.8
Current age (mean)	83.7	84.8	86.4	87.3	88.3	89.2	90.3	91.1	92.0	92.8	93.5

Abbreviations: E4 = ApoE-E4 allele

Table 16. Percentage of participants developing dementia by number of languages spoken and study transition period.

# of langs. spoken	Outcome	% of participants in language category by transition period[1]										
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1	**Dementia**	5.7	3.7	2.9	7.6	10.2	4.8	0.0	3.4	0.0	13.6	6.3
	No dementia	94.3	96.3	97.1	92.4	89.8	95.2	100.0	96.5	100.0	86.4	93.7
2	**Dementia**	9.4	4.8	10.3	6.3	8.6	10.0	4.2	10.0	11.4	4.2	5.3
	No dementia	90.6	95.2	89.7	93.7	91.4	90.0	95.8	90.0	88.6	95.8	94.7
3	**Dementia**	2.0	8.9	13.9	3.4	4.2	0.0	15.4	0.0	0.0	0.0	11.1
	No dementia	98.0	91.1	86.1	96.5	95.8	100.0	84.6	100.0	100.0	100.0	88.9
4+	**Dementia**	0.0	0.0	0.0	14.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	No dementia	100.0	100.0	100.0	85.7	100.0	100.0	100.0	100.0	100.0	100.0	100.0

[1]Participants who died in each transition period are not included in the proportions illustrated by this table.

Abbreviations: langs. = languages

6.2.1.2 Discrete-Time Survival Analysis

In the first part of this analysis, three logistic regression models were developed to analyse the separate, unadjusted effects of each of the following three variables on dementia hazard probabilities over each study transition period (time period between cognitive assessments): multilingualism, ApoE-E4 status, and age at baseline cognitive assessment. ApoE-E4 and age at baseline cognitive assessment were chosen for assessment in their own crude logistic regression models since these two covariates were found to be significantly associated with dementia status in the descriptive analyses. Multilingualism was categorized into four levels, while ApoE-E4 status was categorized according to the absence or presence of an E4 allele. Age at baseline cognitive assessment was categorized into three levels: participants aged less than 80 years at baseline, participants aged 80 to <85 years at baseline, and participants aged 85 years or older at baseline. Values for all hazard probability estimates (the conditional probability that an individual developed dementia in the stated time period, given that they were at risk) generated by regression models can be found in the tables included in Appendix I.

The results of the three unadjusted models are displayed individually in Figures 10, 11, and 12. The results of all three unadjusted models are displayed together, for the sake of comparison, in Figure 14. According to Figure 10, participants speaking two or three languages had higher estimated hazard functions (chronological patterns of conditional hazard probabilities) of dementia development than participants speaking one language (the reference group); however, these differences between hazard function estimates were not statistically significant. Participants who spoke two languages exhibited an estimated dementia hazard function that was approximately 56% higher than the hazard function for participants who

spoke one language (p=0.06). Participants who spoke three languages had an estimated dementia hazard function approximately 13% higher (p=0.71) than the hazard function associated with the reference group. Conversely, the estimated hazard function associated with the participants who spoke four or more languages was approximately 16% lower than the estimated hazard function associated with participants who spoke one language. This reduction, however, was not statistically significant (p=0.08).

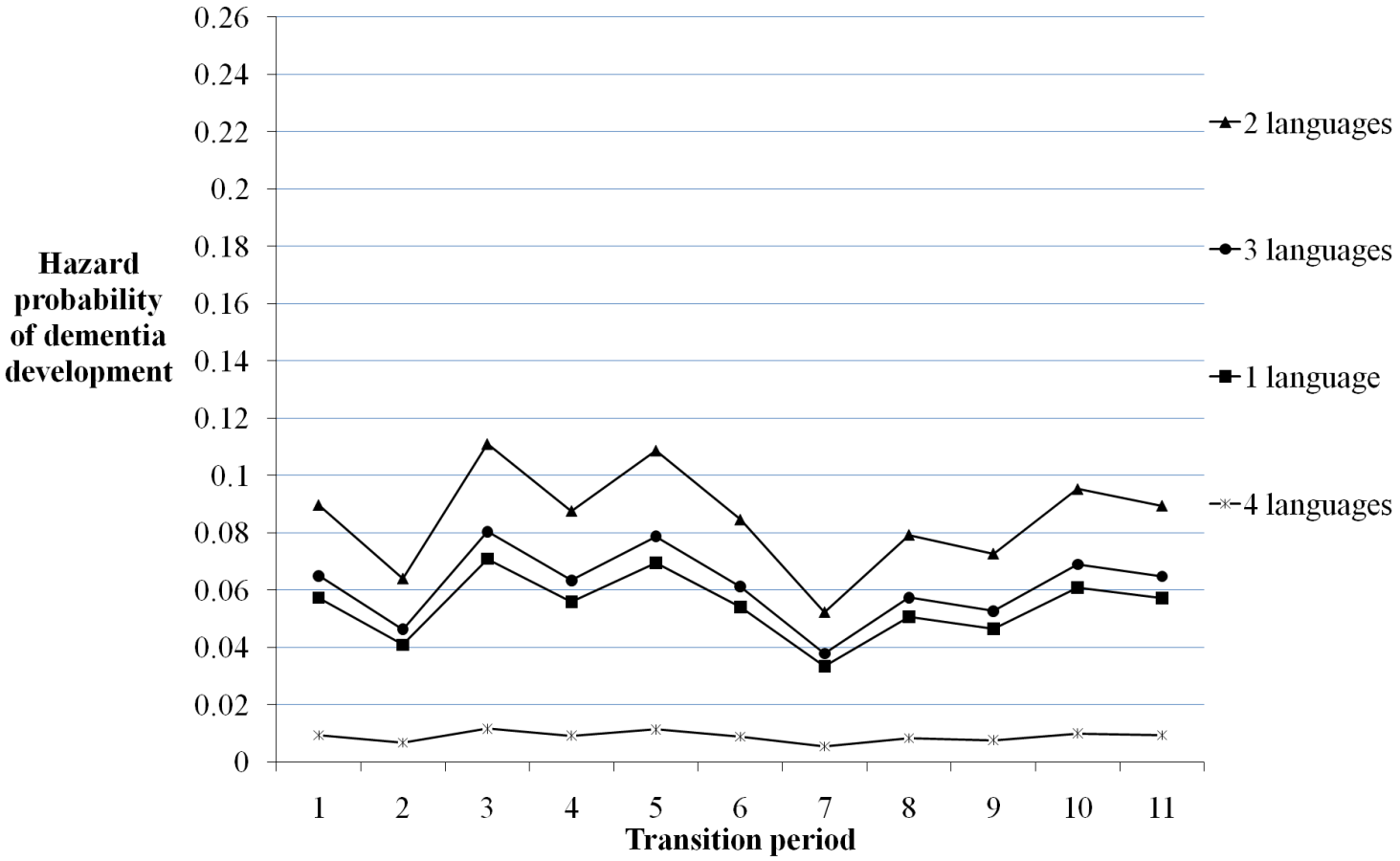

Figure 10. Estimated hazard functions for dementia development by multilingualism (four-level definition, where participants who spoke one language constituted the reference group). Hazard functions and probabilities were generated using a model that included only multilingualism and transition period.

Participants who possessed an ApoE-E4 allele had a significantly higher estimated dementia hazard function than the hazard function attributed to participants who possessed no ApoE-E4 alleles (by 2.18 times; $p<0.0001$). These functions are illustrated in Figure 11. According to the model assessing the crude effects of baseline age category (Figure 12), participants older than 85 years at baseline cognitive assessment had a hazard function significantly higher (4.2 times higher; $p<0.0001$) than the hazard function estimate attributed to participants younger than 80 years at baseline assessment. Similarly, participants between the ages of 80 and 85 at baseline had an estimated dementia hazard function 1.6 times higher than the participants in the lowest baseline age category ($p=0.05$).

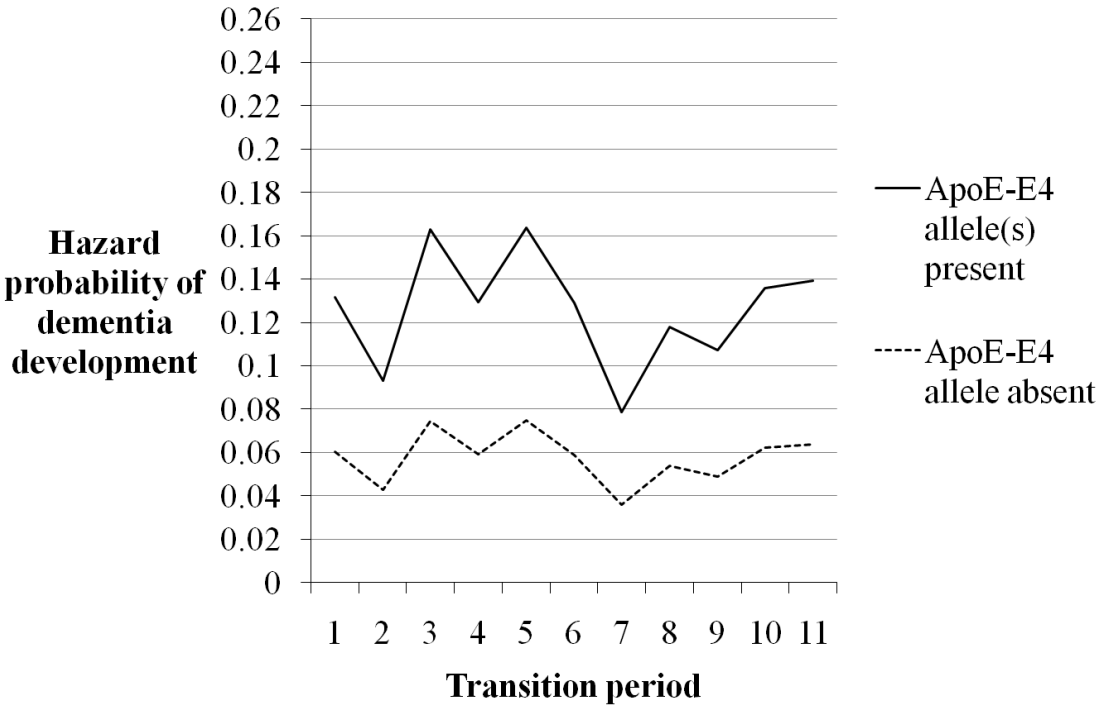

Figure 11. Estimated hazard functions for dementia development by ApoE-E4 status. Hazard functions and probabilities were generated from a model that included only ApoE-E4 status and transition period. ApoE-E4 = Apolipoprotein E-E4 allele

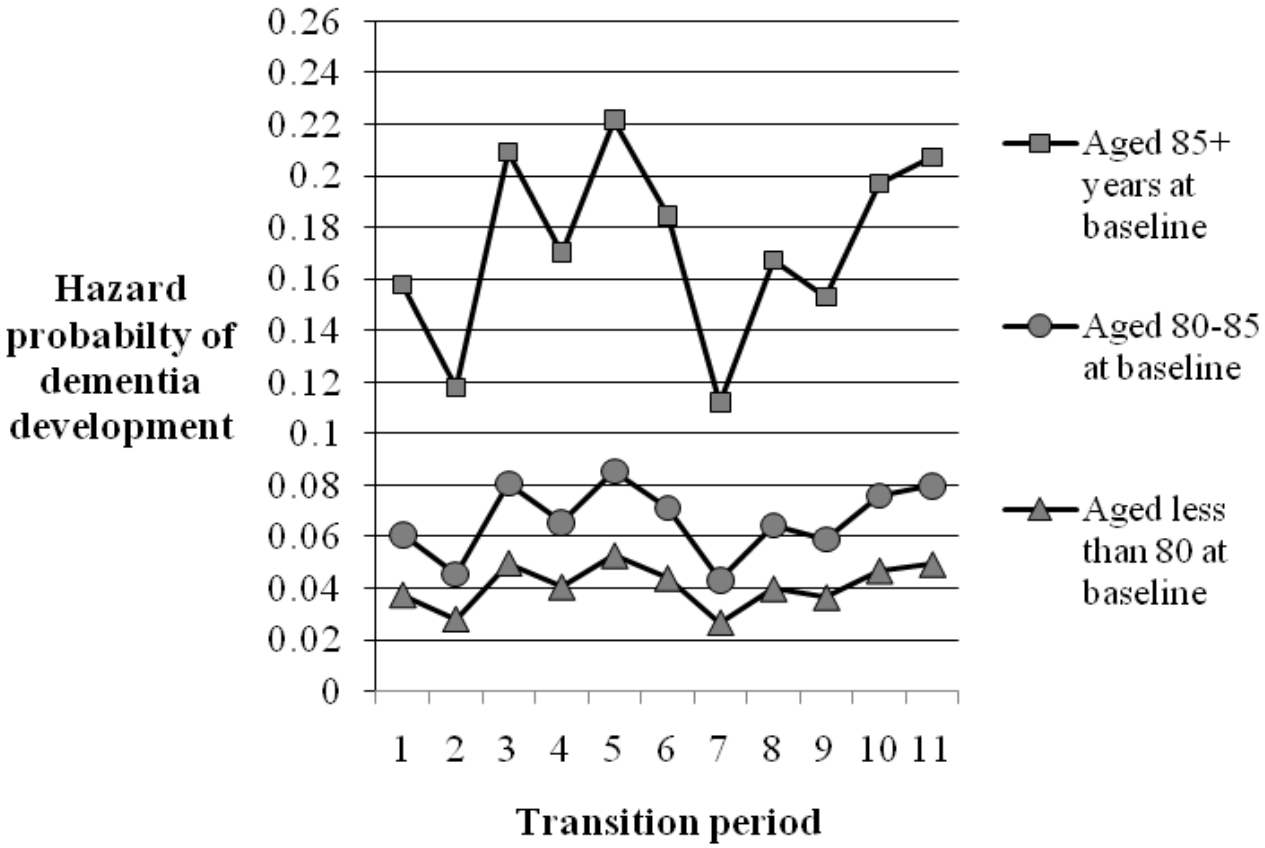

Figure 12. Estimated hazard functions for dementia development by age at baseline cognitive assessment. Hazard functions and probabilities were generated using a model that included age category and transition period.

Figure 13 presents the dementia hazard function estimates produced by each of the previous three models together, in order to facilitate the comparison of their respective sizes. The hazard functions generated by these previous three models all exhibited similar patterns over time. Overall, time did not result in significant net differences in hazard probabilities for dementia development; the estimated dementia hazard functions were constant and persisted throughout the follow-up period of the study. Generally, each estimated hazard function displayed in Figure 13 exhibited a decrease in transition period two, only to increase in transition period three. All hazard function estimates eventually hit their minima in transition period seven, but by the end of the study period (the end of transition period eleven) had returned to heights similar to that from which they began.

A singular model of dementia risk adjusting for all three variables (multilingualism, ApoE-E4 status, and baseline age category), in addition to the effect of time, was also analyzed. Dementia hazard odds ratios produced by this model are displayed in Table 17.

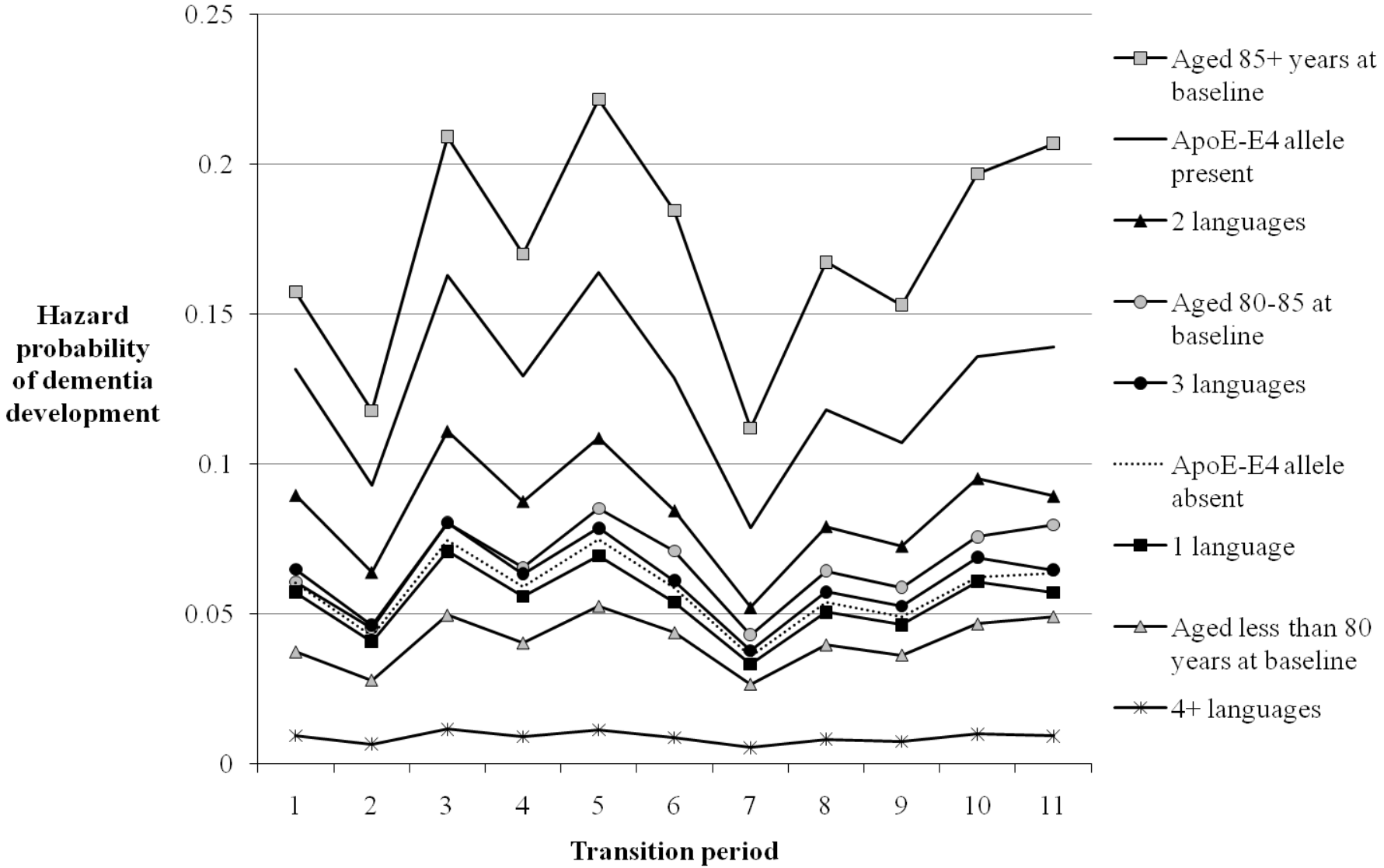

Figure 13. Illustration of all previous three sets of estimated hazard functions combined, for the sake of comparability. Note that the functions displayed were generated not from one model adjusted for all three variables and transition period, but from three different models: a model considering language fluency category and transition period; a model considering ApoE-E4 status and transition period; and a model considering baseline age category and transition period.
Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele

Table 17. Dementia hazard odds ratio estimates generated from a model of dementia hazard probability adjusted for multilingualism, ApoE-E4 status, baseline age, and transition period.

Parameter	OR	95% CI
Multilingualism		
Speaking two languages vs. one	1.56	0.98, 2.55
Speaking three languages vs. one	1.24	0.62, 2.40
Speaking four or more languages vs. one	**0.13**	**0.01, 0.65**
ApoE-E4 status		
Possession of an ApoE-E4 allele	**2.53**	**1.55, 4.05**
Age category[1]		
80 to less than 85 years old	**1.67**	**1.01, 2.76**
85+ years old	**4.80**	**2.90, 8.05**
Transition period[2]		
1	0.58	0.18, 2.61
2	0.44	0.13, 2.03
3	0.83	0.26, 3.71
4	0.70	0.21, 3.21
5	0.93	0.27, 4.26
6	0.79	0.21, 3.83
7	0.49	0.10, 2.65
8	0.76	0.17, 3.91
9	0.70	0.14, 3.78
10	0.94	0.19, 5.11

Bolded estimates indicate statistical significance.
[1] Where the reference group consisted of participants aged less than 80 years.
[2]No estimate available for transition period 11 as the parameter estimate was equal to zero.
Abbreviations: OR = odds ratio estimate; 95% CI = 95% confidence interval; ApoE-E4 = Apolipoprotein E-E4 allele

Judging by the hazard functions generated by the previous models, time appeared to have a relatively limited effect on the hazard probability estimates for dementia, as the dementia hazard function estimates were constant and persisted throughout the follow-up period of the study. Based on this observation, the importance of time and all other covariates of interest (including those not found to be statistically significant in the descriptive analyses, such as educational level, occupation, and immigrant status) in estimating dementia hazard probability were systematically tested. Using the log likelihood test (see Section 5.2.2), it was determined that the most suitable model for dementia probability estimation contained a four-level definition of multilingualism, age at baseline cognitive assessment (either a continuous or three-level categorical definition), and ApoE-E4 status. In order to facilitate graphical interpretation of hazard functions, a model comprised of multilingualism, ApoE-E4 status, and a categorical age at baseline assessment variable was analysed, while a model comprised of multilingualism, ApoE-E4 status and a continuous age at baseline variable was analysed to generate hazard ORs and 95% CIs (see Appendix I, Table 11).

Since these preferred models did not take the effects of time into consideration, the resultant hazard function estimates were linear when graphed over time (i.e., the functions had slopes equal to zero). Therefore, dementia hazard probabilities were interpreted as having one value for a given individual over the entirety of the study period (for all 11 transition periods). For example, according to Figure 14, an individual aged 85 years or older at baseline who had at least one ApoE-E4 allele and who spoke three languages had a hazard probability estimate for dementia development equal to approximately 0.35 for the duration of time they were enrolled in the Nun Study (1-11 years, depending on the given individual). Similarly, an individual aged 85+ years at baseline who possessed an ApoE-E4 allele but spoke four languages had a conditional

hazard probability estimate for dementia development equal to approximately 0.04 for the duration of time they were enrolled in the Nun Study. These estimates are displayed graphically in Figure 14 and displayed in Table 9 of Appendix I.

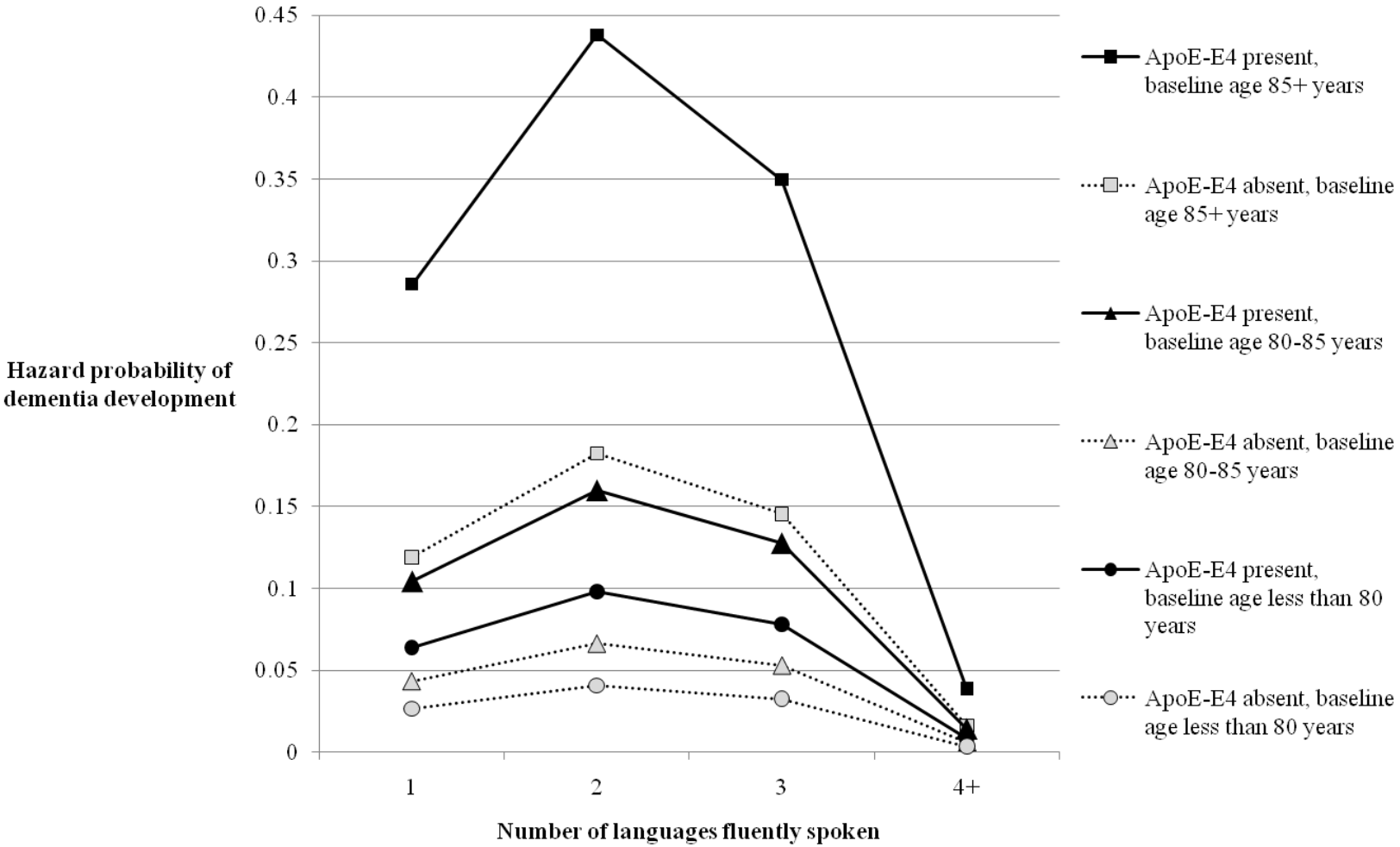

Figure 14. Estimated hazard functions for dementia development, according to category of baseline age, ApoE-E4 status, and multilingualism. Hazard functions and probabilities were estimated using a model adjusted for all three variables.

Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele

According to the functions of dementia hazard probability generated from a model adjusted for multilingualism, ApoE-E4 carrier status, and age at baseline cognitive assessment (Figure 14), speaking 2 or 3 languages was generally associated with a higher dementia risk than speaking one language, if ApoE-E4 status and baseline age were held constant; however, these differences in dementia hazard estimates were not statistically significant. Speaking four or more languages, on the other hand, was associated with a significantly lower dementia risk than speaking one language, given ApoE-E4 status and baseline age were held constant. The risk of dementia was approximately 86% lower ($p=0.05$) for participants speaking four or more languages than participants with similar ages and ApoE-E4 profiles who only spoke one language.

The possession of an ApoE-E4 allele was generally associated with significantly higher dementia risk compared to not having an ApoE-E4 allele, when all other factors were held constant. The dementia hazard probability associated with having an ApoE-E4 allele was 2.4 times higher ($p=0.0003$) than participants without an ApoE-E4 allele. Older age at baseline was also associated with greater dementia risk; participants aged 80 to less than 85 at baseline assessment had dementia risks 1.6 times higher than participants aged less than 80 at baseline ($p=0.05$), given they had similar ApoE-E4 profiles and spoke the same number of languages. Participants aged 85 and above at baseline were 4.5 more likely to develop dementia than participants aged less than 80 years at baseline ($p<0.0001$), given they had similar ApoE-E4 profiles and spoke the same number of languages.

While these trends of dementia risk generally held true, an interaction existed between the number of languages spoken and the other risk factors analysed in this model. Therefore, the protection against dementia conferred by speaking four or more languages could reduce a given

participant's dementia risk, despite the presence of an ApoE-E4 allele or an older baseline age. For instance, a participant who spoke four or more languages and had an ApoE-E4 allele and a baseline age of 85 years or older had a similar risk of dementia (dementia hazard probability = 0.04) as a participant who spoke one language who did not have an ApoE-E4 allele and was younger than 80 years at baseline (dementia hazard probability = 0.03).

6.2.2 Sensitivity Analysis in Linguistic Ability Sub-Sample

The first sensitivity analysis in Research Question 1 analysed the relationship between AD and multilingualism in the context of linguistic ability, and produced different results than when the analyses did not adjust for linguistic ability. Therefore, a sensitivity analysis considering only participants with complete linguistic ability information was performed as part of addressing Research Question 2, in order to see if the relationship between multilingualism and dementia hazard probability would also be different when evaluated in the context of participant linguistic ability. Please refer to Section 4.3.2.2 for a detailed description of the derivation of the sub-sample for this sensitivity analysis.

6.2.2.1 Descriptive Statistics

Table 18 presents the characteristics of this sub-sample by dementia status. There were 106 participants in the total sub-sample, 28 of whom developed dementia (26.4%). Overall, 68.8% of the total sub-sample spoke at least two languages. When split by dementia status, however, it was found that a significantly greater proportion of control participants spoke at least two languages, compared to participants who developed dementia (74.4% versus 53.6%). There were 24 control participants who spoke three or more languages (24/78; 30.8%), compared to just four participants (4/28; 14.3%) who developed dementia. All participants in

this sub-sample had completed at least high school. However, a significantly greater proportion of participants developing dementia had a lower level of education (high school or less) compared to control participants (14.3% of participants with dementia versus 0% of controls). The proportion of participants in each grammatical complexity quartile was similar across dementia status categories. However, this was not the case with idea density: a significantly greater proportion of participants developing dementia had low idea densities (idea density scores in the lowest quartile) compared to control participants (32.1% versus 7.7% of controls). Conversely, a larger proportion of control participants had idea densities in the highest quartile (quartile four) compared to participants who developed dementia (33.3% versus 17.9%, respectively), but this difference was not statistically significant. A significantly greater proportion of participants developing dementia had an ApoE-E4 allele (10/28; 35.7%) compared to control participants (11/78; 14.1%). None of the participants in this sub-sample were immigrants, and most of the sub-sample held teaching occupations (98.1% of the entire sub-sample). Each dementia status group had one participant that was not a teacher by profession. Participants in each dementia status group had similar ages at baseline cognitive assessment, although control participants attended more follow-up cognitive assessments (mean = 7.0 assessments; SD = 3.6) than participants who developed dementia (mean = 4.8 assessments; SD = 3.1).

Table 18. Linguistic ability sub-sample characteristics by dementia status (n=106).

Variable		Total (n=106)	Control[1] (n=78)	Demented[2] (n=28)
Multilingualism				
# of languages (%)	1	31.2	25.6	46.4*
	2	42.4	43.6	39.3
	3	21.7	25.6	10.7
	4	2.8	2.6	3.6
	5	1.9	2.6	0.0
2+ languages (%)		68.8	74.4	53.6*
Covariates				
Age at baseline (years)	mean (SD)	79.8 (2.8)	79.8(2.7)	79.9 (3.2)
# of follow-up assessments	mean (SD)	6.3 (3.6)	7.0 (3.6)	4.8 (3.1)**
Education (%)	Grade school	0.0	0.0	0.0
	High school	3.8	0.0	14.3**
	Bachelor's degree	36.8	38.5	35.7
	Master's degree +	59.4	61.5	50.0
Occupation (%)	Teacher	98.1	98.7	96.4
	House sister	0.0	0.0	0.0
	Other[3]	1.9	1.3	3.6
Immigrant to USA (%)		0.0	0.0	0.0
Possession of an ApoE-E4 allele (%)		19.8	14.1	35.7*
Idea density quartile (%)	1 (low)	14.1	7.7	32.1**
	2	27.4	26.9	28.6
	3	29.3	32.0	21.4
	4 (high)	29.2	33.3	17.9
Grammatical complexity Quartile (%)	1 (low)	15.1	14.1	17.9
	2	27.4	28.2	25.0
	3	30.2	29.5	32.1
	4 (high)	27.4	28.2	25.0

*p <0 .05, **p<0.01, ***p<0.001

[1] Participants did not have dementia (DSM-IV criterion) at any assessment.

[2] Participants were dementia-free at baseline assessment but ultimately demonstrated cognition consistent with dementia (DSM-IV criterion).

[3]An example of another occupation held by participants was a nurse's aide.

Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele; SD = standard deviation; DSM-IV = Diagnostic and Statistical Manual of Mental Disorders, 4th Edition

6.2.2.2 Discrete-Time Survival Analysis

The importance of all covariates of interest, including time parameters (i.e., each of 11 transition periods), to dementia estimates was systematically tested using the log likelihood test (see Section 5.2.2. for additional details concerning this test). The most suitable model for dementia probability estimation considered ApoE-E4 status, idea density, and a three-level education variable (classifying participants as having high school-level education or less, attainment of a bachelor's degree, or attainment of at least a master's degree). Multilingualism was not significantly associated with dementia likelihood, yet it was still included in the model so that its relationship with dementia and the other variables could be understood. ApoE-E4 status, idea density, and educational level were also considered in this analysis. The parameter estimates attributed to the second, third and fourth idea density quartiles were found to be similar, compared to the estimate attributed to the first (lowest) quartile of idea density (see Appendix I, Table 12); therefore, a two-level idea density variable (i.e., first quartile of idea density ranking versus the combination of the second, third, and fourth idea density quartiles) was instead employed by this analysis in order to facilitate the graphical representation of hazard functions (Figure 16). Hazard probability estimates and OR values have been included in Tables 13 and 14 in Appendix I.

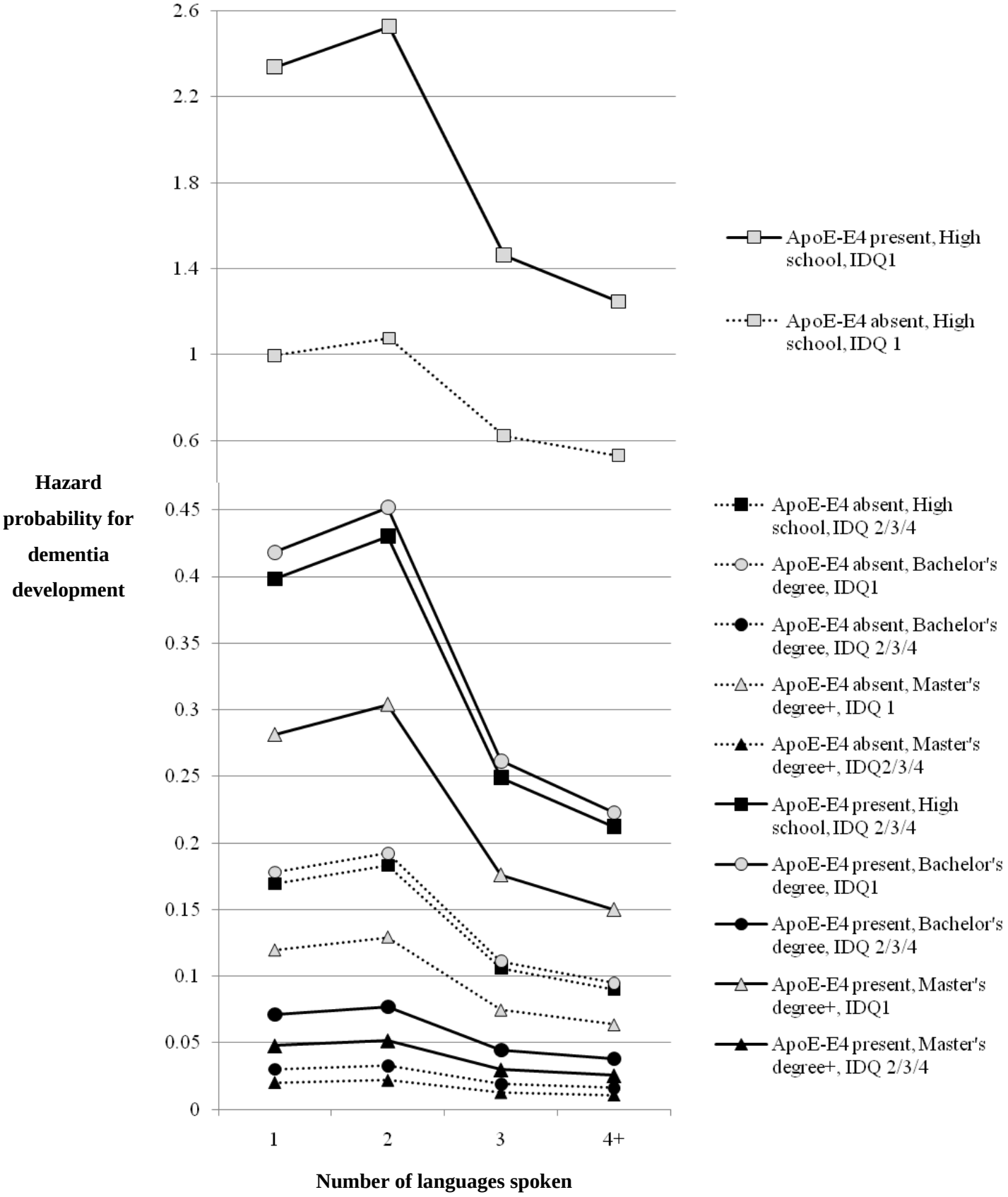

Figure 15. Estimated dementia hazard functions in participants according to ApoE-E4 status, education, idea density quartile, and multilingualism. Functions with dashed lines are those of participants without an ApoE-E4 allele.
Abbreviations: IDQ = idea density quartile; ApoE-E4 = Apolipoprotein E-E4 allele

While speaking a greater number of languages were observed to affect the dementia hazard probability estimates compared to participants with similar ApoE-E4 profiles, educational levels, and idea density scores, none of these differences were statistically significant. According to Figure 15, participants who spoke two languages had higher dementia hazard probabilities than participants who spoke one language (with similar levels of education, ApoE-E4 profiles, and idea density capabilities). On the other hand, participants who proficiently spoke three languages had lower dementia hazard probabilities than participants who spoke one language (when all other variables were held constant). Participants speaking four or more languages once again had the smallest dementia hazard probability estimates compared to all other categories of language proficiency, when all other variables were held constant. When all other participant characteristics were similar, participants with an ApoE-E4 allele had estimated hazard functions 2.35 times higher than the functions relating to participants without an ApoE-E4 allele; however, this difference in hazard function estimates was also not statistically significant.

Participants with higher levels of education had significantly lower dementia hazard functions compared to participants with high school education or less (when all other variables were held constant): individuals with Bachelor's degrees had hazard functions 82% lower than those who had educations no further than high school ($p=0.01$), while participants with Master's degrees or higher had dementia hazard functions 88% lower than those participants with no more than a high school education ($p=0.002$). Participants who scored in the second, third, or fourth quartiles of idea density had significantly lower hazard function estimates compared to participants with idea density scores in the first quartile with similar ApoE-E4 profiles, educational levels, and proficiencies in a similar number of languages (83% lower; $p=0.0001$).

The highest overall dementia hazard function estimate belonged to participants with an ApoE-E4 allele, lowest level of education (high school or less) and idea density scores ranking in the lowest quartile. Participants with these attributes had the highest estimated hazard probabilities, regardless of the number of languages they spoke. The next highest hazard function corresponded to participants without an ApoE-E4 allele, with high school education or less, and idea density scores in the lowest quartile. Participants with high idea density scores (i.e., in any of the top three quartiles) who did not have an ApoE-E4 allele and had at least a Master's degree were the least likely of any participants to develop dementia during the study period. While participants scoring lower in idea density or having less education generally were more susceptible to developing dementia, individuals with these at-risk features were able to decrease their dementia likelihoods if they spoke at least three languages. For example, an ApoE-E4 allele with a high school-level education speaking four languages and a high idea density score had an estimated dementia hazard probability lower than an individual with a Master's degree, a similar ApoE-E4 profile, who had a low idea density and spoke only one language. Similarly, a participant with a more at-risk ApoE-E4 profile could have been less susceptible to dementia if they had attained a high level of education or had high idea density. Overall, the effect of multilingualism appeared to be somewhat diminished compared to analyses that were not adjusted for idea density. It was noticeable, however, that the findings from the present model revealed decreased susceptibility in those speaking three or more languages, whereas the previous model only found a decrease in dementia susceptibility to be associated with speaking four or more languages. Speaking two languages was again associated with an increased dementia likelihood. Similar to multilingualism, the effect of ApoE-E4 status on dementia likelihood was also lessened in this analysis. Idea density and education had the most significant

relationships with dementia likelihood; higher idea density and more education were significantly associated with decreased dementia likelihoods.

7.0 Discussion

7.1 Study Findings

According to the literature, multilinguals are suspected to be more resistant against dementia than monolinguals, due to having higher cognitive reserve, and possibly also brain reserve, levels. Higher levels of reserve are suggested to develop in multilinguals as a result of advantages they experience (over monolinguals) in executive control tasks (Bialystok et al., 2007; Craik et al., 2010). The present investigation found multilingualism to be associated with lower odds of developing AD. However, this relationship was present only when grammatical complexity and ApoE-E4 allele data were incorporated into the analyses. The multilingual participants least likely to develop AD were those without an ApoE-E4 allele (when controlling for grammatical complexity). When time of dementia onset was compared between multilinguals and monolinguals, dementia hazard functions for all participants were constant and persisting throughout the entire study follow-up period, regardless of the number of languages spoken. Multilingualism was associated with decreased dementia risk; however, only speaking four or more languages was associated with significantly decreased dementia risk than speaking one language, when ApoE-E4 profiles and baseline ages were held constant. The association between speaking four or more languages and decreased dementia risk appeared to be stronger than the association between decreased dementia risk and the absence of an ApoE-E4 allele or a lower baseline age; for instance, older participants at baseline with an ApoE-E4 allele who spoke four or more languages had similar dementia risks as younger monolingual participants without an ApoE-E4 allele. When linguistic ability was also held constant across participants, proficiency in three or more languages was related to decreased dementia likelihood. In this analysis, however, the association between multilingualism and dementia risk

was less significant than the associations between dementia and education or dementia and idea density. In all study results, occupation and immigrant status did not appear to be associated with any outcomes, nor did modify the relationship between multilingualism and AD or dementia. Overall, the present results highlight the importance of evaluating multilingualism's relationship with late-life cognition in the context of linguistic ability, ApoE-E4 status (especially when AD is the outcome) and level of education (especially when dementia is the outcome). Some results were inconsistent with those found in previous studies, which was not surprising given our methodologies were novel to the research area. The present study also defined multilingualism according to the number of languages participants reported speaking by means of self-report; no definition of language proficiency was provided by investigators, and proficiency testing was not conducted. Previous studies have required multilingual participants to be "balanced", which relates to regular use of at least two languages for the majority of life (from at least early adulthood onward). Therefore, the multilingualism definition employed in this study was less strict than that used by previous studies of the relationship between multilingualism and AD or dementia. However, our study contributes new evidence on the association of multilingualism with AD and dementia, with methodological strengths in longitudinal data, the breadth of covariates considered, and the absence of common confounders. Contributions from different types of studies (i.e., using various designs and analytic methods) are needed in order to fully understand the relationship between multilingualism and AD or dementia.

7.1.1 Research Question 1

The aim of Research Question 1 was to investigate whether multilingualism was associated with AD risk (where a diagnosis of AD required both clinical dementia and the presence of AD neuropathology). It was surprising that a large proportion of participants in the main analyses spoke more than one language (71.4% of total sample; 71.8% and 70.8% of control and AD groups, respectively), especially since only 4.5% of the total sample were immigrants to the U.S. It is possible, however, that participants developed additional language proficiencies in order to teach while on placements or during their placements abroad (for instance, some participants reported speaking Chamorro, a Malayo-Polynesian language spoken by the people of Guam and other Mariana Islands). According to their official website (accessed May 2011), School Sisters of Notre Dame have a mission to "empower people, especially...women and children, to reach the fullness of their potential" through education. Therefore, many of the participants may have been highly motivated to learn new languages, regardless of whether they taught abroad. The educational mission of participants also helps to explain why 49.0% of the total analytic sample had attained a Master's degree or higher and 93.0% of the analytic sample held teaching occupations. The School Sisters of Notre Dame congregation originated in Bavaria (Germany), and all of its U.S. chapters were began by German immigrants, which may have influenced later generations of School Sisters to develop proficiencies in German or other European languages. Given that 41.4% of participants with language proficiency data spoke German, 18.3% spoke French, 12.6% spoke Spanish, and 10.8% of participants spoke Polish, this may have indeed been the case. Similarly, while the majority of participants were not immigrants themselves, their parents were likely to have been immigrants. Therefore, the participants in the current investigation may have learned additional

languages in childhood from their parents. Whatever the reason, this large proportion of multilingual participants may be another characteristic unique to this special population, and should be considered when interpreting the results.

Regardless of the number of levels used to define the multilingualism categories, multilingualism was not significantly associated with AD risk. According to the four-level definition of multilingualism, participants who spoke four or more languages were less susceptible to AD compared with monolinguals (OR = 0.61; 95% CI = 0.06-6.15). This measure of association was not statistically significant and had a relatively wide confidence interval, presumably due to the relatively small number of participants proficient in four or more languages in the sample. Nevertheless, this finding was interesting since this group of participants was the only multilingual group with decreased AD risk in this analysis, compared to monolinguals. The only factor significantly associated with AD risk was ApoE-E4 status; a significantly larger proportion of participants developing AD (50.8%) possessed an ApoE-E4 allele compared to control participants (8.7%) and the odds of AD in participants with an ApoE-E4 allele were approximately 12 times higher than non-carriers when considering both the four-level and two-level definitions of multilingualism.

ApoE-E4 allele carriers were expected to be more likely to develop AD, given the body of evidence concerning the influence of the ApoE-E4 allele on sporadic AD risk. However, the present analyses were the first to evaluate multilingual ability in the context of ApoE-E4 status. Based on the present analyses, genetics appeared to have a greater association with AD risk than multilingualism (or any other variable in these analyses). An interaction between ApoE-E4 status and multilingualism was also not significant, indicating the association between multilingualism and AD risk did not vary by ApoE-E4 status.

Other Nun Study publications have demonstrated the importance of grammatical complexity with respect to cognitive reserve and AD risk (Riley, Snowdon, Desrosiers, & Markesbery, 2005; Snowdon et al., 1996; Snowdon, Greiner, & Markesbery, 2000; Tyas et al., 2009). While the ability to speak multiple languages is hypothesized to bestow certain cognitive advantages to multilingual individuals, multilinguals may also have less developed vocabularies in each of their known languages than monolinguals (Bialystok et al., 2008; Craik & Bialystok, 2005; Rosselli et al., 2000). Thus, controlling for differences in linguistic ability (such as vocabulary size or related measures, like grammatical complexity) is important to the study of the association between multilingualism and AD (or other late-life cognitive outcomes) (Riley et al., 2005).

When grammatical complexity and idea density were included in the linguistic ability sensitivity analysis for Research Question 1, the association between multilingualism and AD risk became more apparent. Furthermore, keeping grammatical complexity constant across all participants also showed that ApoE-E4 status was again strongly associated with AD risk. In fact, when the distribution of ApoE-E4 genotype was examined within groups of AD and control participants, the dose-response relationship between increasing number of ApoE-E4 alleles and AD risk was especially clear. This means that, without controlling for linguistic ability, this association and its magnitude would have gone unnoticed. While logistic regression models were unable to reliably produce estimates within the linguistic ability sub-sample, a strong relationship between multilingualism and AD was noticeable when numbers of languages spoken were compared between participants with and without AD (Figure 8). A larger proportion of participants who developed AD were monolingual. Furthermore, a larger proportion of control participants were proficient in multiple languages. While these differences

in proportions were not statistically significant, it was clear that participants with and without AD differed according to the number of languages spoken. These observations gave reason to investigate whether the relationship between multilingualism and AD risk differed according to ApoE-E4 status, especially given the demonstrated association between ApoE-E4 and AD in this sample. Multilingualism did not appear to be associated with AD among participants possessing an ApoE-E4 allele (after controlling for age at last cognitive assessment). Among participants without an ApoE-E4 allele, however, multilingualism appeared to be associated with protection against AD as only monolinguals developed AD. This finding suggests that, in participants without the ApoE-E4 allele risk factor, the ability to speak multiple languages may reduce the likelihood of AD. In participants with an ApoE-E4 genetic risk factor, on the other hand, the reduction in AD likelihood associated with multilingualism may be outweighed by genetic predisposition (from ApoE-E4 allele possession). Since multilingualism has been hypothesized to protect against cognitive decline through enhancing cognitive reserve, it is possible that this protective effect is not as robust for cases where neuropathology accompanies cognitive decline (as in AD cases, as shown by Research Question 1), compared to cases of dementia (as demonstrated by the results of Research Question 2). Whether the association between decreased AD and dementia risk and multilingualism results from an increase in cognitive reserve due to heightened executive control, another multilingual advantage in cognitive processing, or another factor associated with multilingualism altogether, remains to be clarified by future studies able to control for both linguistic ability and ApoE-E4 status.

When multilingualism was categorized into two levels and analysed in the context of grammatical complexity and idea density, ApoE-E4 status was not found to significantly influence AD development. Greater grammatical complexity ability, however, was significantly

related to reduced AD susceptibility. Proficiency with two or more languages was associated with a decrease in AD odds, when grammatical complexity was held constant. This decrease was not statistically significant, although the sample used for this sensitivity analysis was relatively small (n=46) and thus, may have not had adequate statistical power.

Past Nun Study analyses have also found idea density to be more strongly related to both cognitive impairment and AD neuropathology than grammatical complexity (Snowdon et al., 1996), which is interesting given that the present results did not find idea density to be significantly associated with AD, although grammatical complexity was. However, this investigation has been the first study of the Nun Study population to analyse these linguistic ability variables while considering the effect of multilingualism; therefore, the emergence of new relationships between linguistic ability variables and AD is understandable.

A limitation of controlling for linguistic ability within these analyses was that the sample size was made considerably smaller than the sample used in the main analyses, as only participants with complete linguistic ability data were included in the sub-sample. Thus, some associations may have gone undetected due to limited statistical power. The sample utilized in this sensitivity analysis did not contain any participants who were immigrants or non-teachers. All participants had completed at least high school, and participants were significantly younger at last cognitive assessment (and death) than participants excluded from this sub-sample. Therefore, it is possible that the participants included in this sensitivity analysis were different than the rest of the Nun Study population and the general population, which might restrict the reproducibility of the observed association in the future. Furthermore, this feature may also limit how applicable these findings are to the general population. Our findings can still largely contribute to the understanding of AD etiology or risk factors, however, as the disease process

present in our study's participants would not be any different from that present in AD cases in the general population.

The Research Question 1 analyses (main analyses and sensitivity analyses) did not detect a strong relationship between educational level and AD, which was surprising given the established association between decreased AD risk and higher educational level. The present finding might be a result of the relatively high overall level of education present among participants in these analyses. A large proportion of this analytic sample had obtained at least a Bachelor's degree, indicating these participants were highly educated women, especially compared to other women of similar ages. Similarly, it was surprising that occupation was unrelated to AD risk, as career complexity is hypothesized to heighten cognitive reserve and protect against AD (Le Carret et al., 2003). This result was unexpected given the prior assumption that a career as an educator would be more cognitively stimulating than a career as a house sister (or other careers available to non-teaching participants). However, this may have not been the case in this sample. For instance, these participants may have engaged in other, unrecorded, cognitively stimulating recreational activities, which may have compensated for a lack of stimulation during work. Alternatively, these participants may have found their occupations to have the same relative difficulty as the participants with teaching occupations, or had more stimulating social environments.

Career length also proved to be unrelated to AD development, which again was unexpected. Teachers with longer careers were hypothesized to have reduced AD odds, based on the rationalization that longer careers would promote longer cognitive stimulation. This unexpected result may have occurred due to the overall engaging nature of life in a religious order. For instance, a participant's passion for learning and education may not necessarily have

ended when she retired from formal teaching; participants may have engaged in other activities beneficial to cognitive reserve long after retiring from a formal teaching career.

Overall, the results from the Research Question 1 analyses illustrated that multilingualism was not significantly associated with AD, unless data on ApoE-E4 status and/or grammatical complexity were also considered. When these factors were incorporated into the analyses, multilingualism appeared to confer protection against AD, but in only those without the ApoE-E4 genetic risk factor.

7.1.2 Research Question 2

The discrete-time survival analyses aimed to compare the probabilities of dementia development between participants in differing multilingual categories and to describe the differences between the probabilities over time. Existing evidence suggests monolinguals manifest dementia at younger ages than multilingual individuals speaking two languages (Bialystok et al., 2007; Craik et al., 2010) or three and more languages (Chertkow et al., 2010). Based on these findings, monolinguals were hypothesized to exhibit higher dementia hazard probabilities at earlier points during the study period than multilinguals. The results of our analyses, however, only partly agree with our *a priori* hypothesis.

Most typical discrete-time hazard models incorporate predictors associated with each of the discrete-time periods considered by the analyses, as time is usually a significant predictor of the event in question. In the current study, however, dementia hazard function estimates were constant and persisting throughout the follow-up period of the study, regardless of participant attributes. Given that age is the most established risk factor for dementia, a participant's probability of dementia development should theoretically increase over time. While the prevalence of dementia has been suggested to double every six years (Ritchie, Kildea, &

Robine, 1992), results from a meta-analysis of 13 studies examining the relationship between age and dementia suggest that the relative risk of dementia does not continue to accelerate amongst the very old. In fact, the authors suggest that, "dementia may not necessarily be part of the normal ageing process, but is perhaps rather a disease with a maximal lifetime risk between ages 70 and 90, with a possible asymptote over 90 years" (Ritchie et al., 1992). Given the participants in our study had baseline ages varying between 75 and 102 years, dementia risk would not have necessarily increased for every participant over time. Furthermore, it was observed in our sample that a relatively similar proportion of all participants at risk (Table 14) and participants at risk in each category of multilingualism (Table 16) developed dementia during each study transition period. Therefore, the investigation's dementia hazard probabilities were plausible.

Dementia hazard estimates were constant over the entire study follow-up period, meaning time did not significantly alter estimates of dementia risk. Speaking four or more languages, possessing an ApoE-E4 allele, and having a baseline age of 85 or greater did significantly alter dementia risk estimates. When ApoE-E4 status and age at baseline were held constant, participants who spoke two or three languages had dementia risks similar to monolingual participants. This was noteworthy as the ability to speak two or three languages was expected to be associated with a significant decrease in dementia risk compared to speaking one language. On the other hand, participants speaking four or more languages had significantly lower hazard probabilities than monolinguals. These individuals did not, however, experience a delay in dementia onset compared to monolingual participants as the estimated dementia hazard functions were constant over time for all participants regardless of how many languages were spoken. These results were intriguing, as they did not fit with the postulated theory. Why would

the protective association between multilingualism and dementia be present in only the multilinguals able to speak four or more languages? If multilingualism allows cognitive process enhancement and subsequent protection against dementia via heightened cognitive reserve, then why would this protective effect be absent in participants speaking two or three languages?

If the hypothesized mechanism by which multilingualism enhances executive control (and, perhaps, cognitive reserve) is assumed to be correct, then all multilingual participants should have exhibited lower dementia hazard functions compared to monolingual participants. In order to be minimally consistent with the postulated theory of multilingual advantages, participants proficiently speaking two, three, and four or more languages should have exhibited lower hazard function estimates, as well as a dose-response relationship, when compared to the hazard functions of monolingual participants. However, this pattern of estimated hazard functions was not observed among participants with differing language proficiencies. This observation calls the proposed mechanism of dementia protection by multilingual ability into question, and suggests that alternative theories should be considered. While the ability to speak and switch between many languages may bestow advantages in executive control upon a given individual, perhaps these advantages do not necessarily contribute towards cognitive reserve. Alternatively, the personal traits or characteristics of the individuals capable of speaking more than four languages may have been systematically different than those of the rest of the multilingual individuals in this sample, and these traits instead allowed for heightened cognitive reserve. For instance, learning to speak four languages may require an individual to be more highly motivated and passionate about language learning, if one assumes that not all four languages were learned passively during childhood. This drive for learning may also motivate such an individual to master other skills and abilities, which instead might be the source of

heightened cognitive reserve. An individual speaking two languages, on the other hand, may have learned their languages during early childhood and thus, may have not chosen to actively pursue the acquisition of novel languages.

A study of whether individuals speaking multiple unrelated languages (e.g., French and Mandarin) are more likely to have reduced dementia and AD risks than individuals speaking two similar languages (e.g., French and Spanish) might also prove to be an interesting future investigation. Speaking multiple unrelated languages might be considered to be more cognitively challenging than becoming speaking many similar languages. Relating back to the findings of the present study, participants who spoke four or more languages may have been more likely to speak unrelated languages, while participants speaking two languages may have had knowledge of more similar languages. Future investigations will need to evaluate more evidence concerning the relationship between multilingualism and dementia before a hypothesis concerning beneficial language type can be established.

The present findings might also be explained by the concealed influence of another activity unmeasured by this study. For example, Bialystok and DePape (2009) demonstrated that monolinguals who had studied a musical instrument for at least half of their lives had similar executive control advantages over non-musical monolinguals as those experienced by non-musical multilingual participants. Therefore, it is possible that activities unrelated to language proficiency may also allow individuals similar cognitive advantages to those experienced by multilinguals (Bialystok & DePape, 2009). Since the present investigation did not collect information concerning musical ability or other recreational activities of a similar nature, the present analyses were unable to control for these other possible influences. Nevertheless, this possibility presents another explanation for the observed dementia risk

differences found between multilingual individuals (i.e. between participants speaking two, three, and four languages or more) in this study.

Although the present findings are somewhat difficult to connect back to the general theory of multilingualism and enhanced cognitive reserve, they are similar to the results from Chertkow et al. (2010). In their study, Chertkow et al. found that participants who spoke three or more languages were significantly older at onset of dementia symptoms (three or more languages: mean = 78.6 years; SD = 6.0 years) compared to monolingual participants (one language: mean = 76.7 years; SD = 7.8 years). Participants speaking two languages, however, experienced no delay in onset of dementia symptoms (two languages: mean age of onset = 76.7 years; SD = 7.8 years) compared to monolinguals. When the authors restricted their sample to non-immigrant participants (making their sample roughly analogous to the sample used in the present investigation, given the relatively small percentages of immigrant participants in our samples), participants speaking two languages actually experienced dementia symptoms significantly earlier than participants speaking one language (2.6 years earlier). Moreover, participants speaking four or more languages experienced a significant delay in dementia diagnosis when compared to a reference group of participants speaking two or more languages (rather than one language).

Chertkow et al. (2010) compared age at diagnosis between participants who all had developed dementia, while the present investigation calculated conditional probabilities for dementia development using participants with and without dementia. Therefore, our results are not directly comparable to those of Chertkow et al. Nevertheless, it is interesting to note the similarities between some of our results and their findings. Akin to Chertkow et al.'s findings, the present investigation's results from the discrete-time survival analysis also found

participants proficient in two languages to have increased dementia hazard probabilities compared to monolinguals. If our reference group had instead consisted of participants speaking two languages or been combined with monolinguals, participants speaking three or more languages would have exhibited smaller dementia hazard function estimates compared to the reference group.

The aim of Chertkow et al.'s 2010 study was to replicate Bialystok et al.'s (2007) findings. Bialystok et al. compared age at dementia diagnosis between participants speaking at least two languages and monolinguals, and found monolinguals to be younger than multilinguals at the time of dementia diagnosis. Bialystok et al. (2007), however, did not distinguish between participants proficient in more than two languages (i.e., participants were only categorized as speaking one and two or more languages); therefore, similar effects for participants speaking three and four languages may have been present in the sample used by Bialystok et al., but the effects may have been missed as these participants were grouped together with participants proficient in two languages. Chertkow et al. (2010) also suggested that their results may have been explained by differing SES levels between monolinguals and multilinguals, as multilinguals may have lower SES levels than monolinguals and nullifying the late-life cognitive benefit expected in multilinguals. In our investigation, all SES levels were relatively equivalent across participants, yet our results remained similar to those of Chertkow et al.

Other dissimilarities between our study and previous studies with similar aims might also be explained due to the nature of participants examined by our study. For instance, Bialystok et al. (2007), Chertkow et al. (2010), and Craik et al. (2010) all compared the age at dementia diagnosis between multilingual and monolingual patients from memory clinics. While

examinations that make use of data taken from memory clinics are unquestionably important to the advancement of dementia epidemiology, these investigations may miss important aspects of the association between multilingualism and dementia as they do not evaluate information from individuals who remain dementia-free. Since cognitive reserve is truly epitomized by individuals resisting cognitive impairment despite biological predisposition or neurological insult, it is fair to assume that clinic-based studies may not capture all aspects of the relationship between multilingualism and dementia or cognitive reserve. Our study, on the other hand, made use of a discrete-time survival analysis that assessed data from participants who developed dementia as well as from participants who remained dementia-free. Time of dementia onset and the time duration an individual spent free of dementia were also factored into the present study's dementia risk estimates. Our study is the first to evaluate this relationship utilizing longitudinal data from participants with and without dementia; it is also the first to utilize a discrete-time survival analysis to evaluate the association between multilingualism and dementia risk. More evidence derived using different participant populations and study designs is required before any definitive conclusions can be made with respect to multilingualism and dementia risk or the delay of dementia onset.

Regardless of how past multilingual categories were classified and whether they exhibited significant delays in dementia onset, the reduction in dementia risk for participants proficiently speaking four or more languages was undeniable. Moreover, the oldest participants at baseline who had ApoE-E4 alleles (i.e., with the most at-risk biological traits) were estimated to have lower dementia hazard probabilities than other participants if they spoke four or more languages, even if the other participants were less at risk for dementia according to age and ApoE-E4 status but instead spoke one, two, or three languages. For instance, an ApoE-E4

carrier who was 85+ at baseline assessment and spoke four languages was estimated to have a dementia hazard probability of approximately 0.04, which made this individual less likely to develop dementia over the course of the study period as an individual 80 years old at baseline without an ApoE-E4 allele who spoke two languages (probability = 0.66). This reduction in dementia probability for individuals speaking four or more languages despite important genetic and biological risk factors implies that an interaction was present in this analysis. Thus, in this sample, multilingualism may have promoted a greater resistance against dementia manifestation despite effects of other risk factors. Both ApoE-E4 status and age undoubtedly have meaningful influences on dementia likelihood: their associations with dementia risk have been demonstrated in the literature, as well as in the present study. Therefore, the finding that speaking four or more languages may reduce the risk of dementia in this sample, despite the influence of genetics and age, is appreciable even if the exact mechanism remains uncertain.

When dimensions of linguistic ability (i.e., grammatical complexity and idea density) were evaluated along with other covariates in the Research Question 2 linguistic ability sensitivity analysis, multilingualism was not significantly associated with dementia hazard probability. While this lack of statistical significance may have resulted from a smaller sample size than that used in the primary survival analysis (n=106, compared to n=350 in the primary survival analysis), the diminished effects of multilingualism were also displayed by the hazard functions in Figure 15. Nevertheless, this analysis revealed some interesting aspects of multilingualism's association with dementia risk. For instance, participants who spoke two languages had increased dementia hazard estimates compared to participants who spoke only one language. This finding was similar to that of the previous survival analysis, although these differences were both not statistically significant; thus, the risk of dementia associated with

speaking two languages cannot be interpreted any differently than the risk of dementia associated with speaking one language. Individuals proficient in three languages, on the other hand, had reduced dementia hazard probabilities compared to participants proficient in one language, when prior analyses had findings suggestive of the opposite effect (proficiency in three languages had been associated with elevated AD and dementia likelihood, according to Figure 14). While the associations between speaking three languages and dementia risk were not significantly different than the associations between dementia risk and speaking one language in both of the previously mentioned analyses, these results cannot be interpreted as being different from one another. Nevertheless, this observed result illustrates the importance of controlling for linguistic ability when evaluating the association between multilingualism and dementia.

Participants with proficiency in four or more languages had reduced hazard probability estimates; these participants had the lowest dementia hazard probability estimates compared to all other participants speaking an alternative number of languages. Therefore, participants proficient in four or more languages were consistently found to have reduced dementia hazard probabilities across all Research Question 2 analyses. The results of this sensitivity analysis again followed a pattern similar to that described by Chertkow et al. (2010), although this pattern did not meet statistical significance. Therefore, if this pattern was to be replicated and is found to be statistically significant, one would have to critically deliberate why participants proficiently speaking three or more languages would experience any cognitive benefit over participants proficiently speaking only two languages, given both are hypothesized to experience executive function advantages.

Based on the literature, multilinguals exhibit less developed vocabularies in each of their known languages, compared to the vocabularies of monolinguals. Therefore, incorporating idea density in the discrete-time survival analysis was expected to change the observed association between multilingualism and dementia. Analyses unable to control for linguistic ability, on the other hand, would be less able to discern the exact relationship between multilingualism and dementia (or AD). Given that low idea density has been closely associated with significantly greater dementia risk (Riley et al., 2005; Snowdon et al., 1996), investigations without controlling for elements of linguistic ability may actually be more inclined to find multilingualism associated with higher dementia risk. Therefore, for a clearer understanding of multilingualism and dementia's association to be developed, linguistic ability should always be an important variable for analytic consideration.

While multilingualism and ApoE-E4 status were not significantly associated with dementia likelihood, idea density was found to be significantly associated with estimating the probability of dementia development. While this aspect of linguistic ability was associated with dementia risk over the study period, grammatical complexity (the other measure of linguistic ability available for analytic consideration) was not significantly related to dementia development. These findings are in agreement with those of past Nun Study publications (Riley et al., 2005; Snowdon et al., 1996), where idea density was found to have a more consistent association with poor cognition and dementia status, compared with grammatical complexity. Idea density was observed to have the strongest relationship with dementia likelihood of any variable considered in this sensitivity analysis; therefore, these findings suggest that future assessments of the association between multilingualism and dementia should be careful to control for linguistic ability differences.

This sensitivity analysis revealed that participants with higher levels of education had significantly decreased dementia risks. These results also demonstrated educational level to have a dose-response effect with dementia probability: participants with the lowest level of education (high school or less) had the highest dementia hazard function estimates, while participants with the highest level of education (Master's degree or higher) had the lowest dementia hazard functions. Although this difference existed between the highest and lowest education levels, overall the dementia likelihoods associated with Master's and Bachelor's level education were generally similar. This relationship between education and dementia likelihood is consistent with previous findings concerning the association between educational level and dementia, both in the Nun Study (Mortimer et al., 2003; Snowdon, 1997; Tyas et al., 2007) and other populations (Evans et al., 1993; Fratiglioni et al., 1997; Lindsay et al., 2002; White et al., 1994). The current findings are also in accordance with the literature in terms of education having a stronger relationship with dementia risk than AD risk (Mortimer et al., 2003), as education was not significantly associated with AD in any of the current investigation's other analyses. Given that cognitive reserve concerns the efficient utilization of neural structures, while brain reserve concerns the quantities and sizes of neural structures, one possible explanation could be that education is more associated with cognitive reserve than brain reserve. Conversely, biologically-based variables such as ApoE-E4 may be more related to brain reserve and outcomes incorporating an element of neuropathology (such as AD). If education and advanced learning can be rationalized to be more advantageous to cognitive reserve, rather than brain reserve, then this more robust relationship between education and dementia (compared to education and AD) might be explained. The results from this sensitivity analysis also illustrated that participants with lower educational levels were able to reduce their

likelihood of dementia if they had high idea density (Figure 15). High idea density has also been hypothesized to be related to a high level of cognitive reserve; therefore, an additive effect between education and idea density in determining dementia risk would be expected, given they are likely to affect dementia risk by means of similar mechanisms.

The findings of the present sensitivity analysis also provide an example of how the association between genetic influence and dementia risk can be altered, depending on an individual's life experiences or cognitive reserve level. For instance, the possession of an ApoE-E4 allele did not necessarily ensure that a participant had a higher dementia hazard probability estimate than a participant without an ApoE-E4 allele; a participant's idea density score or educational level was more indicative of their estimated dementia risk than the presence of an ApoE-E4 allele. Furthermore, these findings also illustrate how the genetic determination of risk is not necessarily absolute for certain cases of dementia (i.e., cases of dementia resulting from sporadic AD). Similarly, these findings provide an illustration of how multiple risk factors, both biological and experiential, contribute to the risk of dementia in any given individual.

7.2 Study Limitations

With respect to the methodology employed by this investigation, there are certain limiting aspects that should be acknowledged. The following sections address several such limitations. The implications of these limitations should be considered when interpreting the results of this study.

7.2.1 Ascertainment of Multilingualism

Since the number of languages spoken was ascertained by means of self-report questionnaire, and no criteria or testing of multilingual ability was provided by investigators, the accuracy of the exposure measurement was dependent on each participant's interpretation of "language proficiency". Consequently, differential misclassification potentially existed during exposure measurement, as some participants may have held themselves to stricter definitions of language proficiency than others. Aside from potential misclassification, self-reported multilingualism status could be more prone to biases than if an objective method of measurement had been utilized. For example, if a participant was given the impression or had a belief that multilingual ability was preferable to being monolingual, it is possible that they could have exaggerated their reported language proficiency abilities. Conversely, it has been observed that older multilinguals "tend to underestimate their own proficiency levels in both languages [and] as such, individuals who do not report themselves to be [multi]linguals may nevertheless turn out to be [multi]lingual by objective standards" (Mindt et al., 2008; p. 262). Therefore, reporting errors may have occurred in exposure reporting for monolinguals and multilinguals alike. If an underestimation of multilingual participants was to occur in a given study, then the likelihood of finding a significant difference in dementia likelihood would be falsely diminished (results would be biased towards the null). If this was the case in our study,

then some of our null findings might be explained, given some findings suggested that multilinguals experienced no significant advantages in AD avoidance, compared to monolinguals.

Other accounts of multilingual cognitive advantages have measured language proficiency in a more objective manner, whether through fluency testing or through self-report using the language(s) in which they claim to be proficient. Given that this investigation made use of secondary data collected from a questionnaire administered several years before the inception of the Nun Study, this investigation did not have the option to objectively measure participant language proficiency. This questionnaire, however, was administered to convent members with the intent to facilitate future placements, including teaching positions in foreign countries or teaching foreign languages in the United States. It can thus be inferred that a given participant would likely have reported proficiency in only languages in which they believed they could teach, which would require a high degree of language aptitude. Additionally, because this questionnaire was administered years before the Nun Study began, there would have been no reason for participants to suspect that these data would be used to judge their cognitive abilities or reserve capacities, and therefore alter their self-reported language proficiencies. Future studies of multilingualism and late-life cognitive decline should ascertain multilingualism as objectively as possible, using standardized measures (when applicable).

Another limitation of the investigation's exposure measure was that multilingual participants were not established to be "balanced" in their language usage. Balanced multilinguals are defined as individuals who proficiently utilize multiple languages for equal amounts of time every day. Many investigations of potential cognitive advantages associated with multilingualism have required multilinguals to be balanced, since it has been suggested

that multilingual cognitive advantages might manifest only in those utilizing their multiple known languages equally (Bialystok & Majumder, 1998). This hypothesis comes from the idea that multilingual advantages are obvious especially during focused attention and task-switching exercises (Bialystok et al., 2004; Bialystok, 2009; Craik & Bialystok, 2005). These abilities are hypothesized to be strongest in individuals who have had maximal practice at ignoring one language while speaking another – in other words, individuals who switch between languages daily. As this study did not inquire about the frequency or duration of daily language use, it was not determined whether multilingual participants were balanced. While future investigations may choose to define exposure based on this criterion, the practicality of restricting studies to only balanced multilinguals has also been questioned (Bochner, 1996; Chertkow et al., 2010; Segalowitz, 1981; Segalowitz, 1990). For instance, Bochner (1996) noted that individuals who meet the lay definition of multilingual (individuals who use two or more languages in which they are equally proficient) are relatively rare. Bochner (1996) also went on to say that this strict multilingual definition would also exclude individuals who may not be completely proficient in a second language, but may still use a second language as an important mode of communication. Segalowitz (1986) used the term "fluent bilingual" to describe people who can express most ideas equally well in each language and who have decent mechanical fluency in each language but may not meet the definition of a "balanced" bilingual as they may not utilize each of their languages equally every day. This definition may be a more realistic requirement for multilinguals in future investigations. A more relaxed definition of multilingualism may also improve the generalizability of results, as participants meeting a strict definition of multilingualism may be rare and thus different than multilinguals in the general population. On the other hand, a more stringent definition of multilingualism might be more likely to show an

association with AD or dementia. For example, if we had employed a stricter multilingualism definition in our study, the participants who reported speaking four or more languages would likely have still been classified as multilingual, while participants classified as speaking two languages would have been less likely to meet the multilingualism criterion. If this had been the case, then our results may have been more similar to those of past findings. Contributions from studies using varied designs and multilingualism definitions have value in progressing the overall understanding of the association between multilingualism and dementia or AD.

7.2.2 Covariates not Assessed by this Study

While this investigation was able to consider many covariates in its analyses, the use of secondary data also limited the measurement of other potentially important covariates. For instance, a higher level of engagement in mentally stimulating activities (e.g., playing chess, or participating in volunteer activities) has been hypothesized to be associated with a lower AD risk (Fratiglioni et al., 2004; Scarmeas & Stern, 2004; Staff et al., 2004; Valenzuela & Sachdev, 2006). Even though speaking many languages is perceived to be a routine mental exercise, it is still unclear whether other cognitively challenging activities may also work by similar mechanisms. Moreover, it is still unknown whether other mentally stimulating activities could interact or modify the hypothesized cognitive effects of multilingualism. It is also possible that other activities (e.g., routinely playing a musical instrument [Bialystok & DePape, 2009]) may allow for similar cognitive benefits as multilingualism or for a greater level of cognitive reserve to develop than through multilingualism (even though multilingualism might be protective, to a degree, against AD or dementia). Similarly, the magnitude of the positive effects associated with social engagement remains unclear with respect to cognitive reserve. Social interaction may have also altered the observed association between multilingualism and dementia or AD in

our study population, as our participants may have reached their maximum possible cognitive reserve benefit from a source of social engagement. At the current time, not enough is known about different mentally stimulating activities to feasibly propose one to be more valuable to cognitive reserve than another; however, in order to advance this research area, the analysis of data on many mentally engaging activities and covariates is necessary in order to measure how they may (or may not) contribute to cognitive reserve.

The exposure definition utilized in this investigation also did not consider the age at which a given participant acquired a second language. According to the literature, the age at which an additional language is "acquired" (age at which an additional language is fully learned) may influence additional language learning (Hakuta, Bialystok, & Wiley, 2003; Paradis, 2005). Moreover, age of additional language acquisition has also been suggested to exhibit an inverse association with parietal cortex grey matter density in healthy individuals (Mechelli et al., 2004). Data on age of additional language acquisition were not available in this study, as data concerning language proficiency were collected in 1983. At this time, participants would have been at least age 65, since participants were 75+ years in age when the study began in 1991-93. Thus, it is possible that multilingual participants could have learned additional languages at any point in their lives up to this time. If age of language acquisition information is collected as a part of future investigations, it would be interesting to compare AD and dementia outcomes between multilinguals who acquired additional languages in childhood with multilinguals who acquired additional languages in adulthood. While individuals acquiring additional languages in childhood may achieve more developed language proficiencies than individuals learning languages later in life (Paradis, 2005), individuals who meet the challenge

of mastering a new language at a later age might also reap the benefits of a later-life cognitive challenge.

7.2.3 Reduced Sample Sizes due to the Exclusion of Participants

Since the association of multilingualism with AD and dementia was not the primary study question when our data were originally collected, some participants were missing data on variables of interest. Participants missing data on primary variables of interest (e.g., multilingualism or AD/dementia) were excluded. While most of the analyses in this investigation were based on samples that were comparable in size to those used in the literature, analytic power would have been greater if a larger proportion of the original Nun Study population had been retained. Furthermore, response bias may have been an issue in our study, as participants without data on certain variables were not included in the analyses.

While the exclusion of participants due to missing multilingualism data was necessary and no participants lacked information on clinical dementia, the use of the NIA-RI criterion for gauging AD neuropathology in Research Question 1 was more constraining on sample sizes than if this investigation had employed another neuropathological criterion. This was because several participants did not fit into the distinct categories of neuropathology described by the NIA-RI criterion. The categories defined by the NIA-RI criterion assume plaques and tangles are directly correlated; that is, that brains with lesser distributions and/or severities of plaques also have similarly lower tangle distributions (The National Institute on Aging - Reagan Institute (NIA-RI) Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease, 1997). This is not always true, however, meaning some individuals are left unclassified by the NIA-RI criterion and, thus, excluded by this investigation. This limitation has been acknowledged (Nelson et al., 2010) and was anticipated

when the NIA-RI criterion was chosen for defining AD neuropathology in this study. It was originally thought that, if defining the outcomes using the "high likelihood" (for AD cases) and "low likelihood" (for the control participants) NIA-RI categories was too restrictive, participants meeting the criterion for "intermediate likelihood" could also be incorporated into the case group. After exploring this possibility, however, it was apparent that including these participants into the analytic sample made the sample substantially different from excluded participants. Therefore, the potential benefit of adding 40 more participants to the analytic sample was judged to be outweighed by the costs associated with having an analytic sample too dissimilar from the excluded participants.

In spite of this limitation, the NIA-RI neuropathological criterion was still regarded to be the most suitable neuropathological criterion for the outcomes in this study as it considers both of the neuropathological hallmarks of AD (plaques and tangles). There currently exists supporting evidence for the etiological roles of both plaques and tangles in AD. Therefore, both neuropathologies should be considered when evaluating AD neuropathology, and currently the NIA-RI criterion is the most conventional standard for doing so.

The limitations resulting from a small sample size were especially noticeable in the linguistic ability sensitivity analysis conducted in Research Question 1. Only 180 participants in the original Nun Study population had handwritten autobiographies evaluated for linguistic ability (grammatical complexity and idea density). After restricting the sample to include only participants with full data across all variables, 46 participants remained. Since the consideration of linguistic ability appeared to significantly influence the relationship between AD and multilingualism, our study would have been strengthened if more participants had provided linguistic ability data. It is important to note, however, that the Nun Study is rare in having the

capacity to evaluate AD and dementia development in the context of linguistic ability. This is because data were collected using handwritten autobiographies composed by participants at the time they entered the religious order (between the ages of 18 and 22), which were available through access to the convent archives. All participants were blinded to potential study aims at the time at which these autobiographies were written, which means that these data were free of certain response biases present in unblinded studies.

7.2.4 Differences between Analytic Samples and Excluded Participants

While the characteristics of the participants included by the Research Question 1 main analytic sample were generally similar to those of the participants excluded from the analyses (Appendix C, Table 1), the sub-samples used in each of the sensitivity analyses were more dissimilar from the excluded participants. For instance, the participants analysed in the teachers-only sensitivity analysis had a significantly lower proportion of participants from the lesser-educated categories as compared to the participants excluded from this analysis (Appendix C, Table 4). While this suggests that the results from this analysis may not be representative of the entire Nun Study population, this difference is understandable given that participants occupying positions other than teachers were almost entirely from the lowest educational category (10 of 11 participants who occupied positions other than teachers had only completed grade school). Therefore, participants occupying house sister positions may have been assigned these roles due to lower educational attainment; alternatively, house sisters may not have pursued further education as it was not required in their house sister role.

Participants analysed as a part of the linguistic ability sensitivity analysis (Research Question 1) were also significantly different from excluded participants in some respects. For

instance, participants analysed in this sensitivity analysis were significantly younger at age of last cognitive assessment than participants excluded from this analysis. This was due to the fact that the cohort of participants who had provided handwritten autobiographies were younger (mean age = 87.5 years; SD = 5.3) than the rest of the sample without autobiographies (mean age = 90.2 years; SD = 5.7). Participants included in this sensitivity analysis were also all non-immigrants and teachers. This was anticipated, as few participants in the analytic sample held alternative occupations: of all participants with linguistic ability information, only 10 of 180 (5.5%) participants held occupations other than teaching, while two participants were missing occupational data.

Although there were few differences between analytic samples and excluded participants, it was possible that participants in the analyses were systematically different from the rest of the study population across some domains. Thus, if all participants had provided full data on all covariates of interest, the findings of the present investigation may have been different.

7.2.5 Generalizability

The Nun Study population provided a unique opportunity to study associations between multilingualism and AD, as study participants were relatively free of several common confounders present in the general population. Although this was an advantage, it was also a limitation, as results from these analyses may not be entirely generalizable to the rest of the public. Since this investigation employed a study population composed entirely of nuns, the results may be only applicable to women. Although AD risk is not suspected to vary substantially according to gender (Gao et al., 1998; Swanwick & Lawlor, 1999), there may have been other influences (e.g., effects of female hormones or post-menopausal conditions) relevant

to this investigation's results that should be considered before applying these results to the general population.

7.3 Study Strengths

7.3.1 Uniform Study Population Lacking Common Confounders

While the homogeneity of the study sample was a limitation in terms of generalizability to the general public, it was also a large methodological strength as it eliminated or reduced many potential confounders. For instance, study participants had similar reproductive and marital histories. Furthermore, the relationship of interest could be assessed without needing to control for influences of alcohol or tobacco use, as all study participants had similar lifestyle habits. Most participants were highly educated, which is relevant to studying AD in the present time as having a post-secondary education is becoming more common. The participants in this study also had similar adulthood environments: this included access to medical care, social supports, and incomes. Similar incomes across all participants allowed the advantage of controlling for SES-related confounding. Since SES has been suggested to significantly alter the association between multilingualism and cognitive benefits (and consequently, late-life cognitions such as dementia) (Chertkow et al., 2010; Mindt et al., 2008; Morton & Harper, 2007; Morton & Harper, 2009), controlling for SES was a large asset to this investigation. Other studies in this research area have been criticized for their inability to assess SES as a confounder, either due to lack of data or because of the difficulties separating influences from immigration and SES. These two covariates are often intertwined within the general population, and it is often difficult to tease out the effects of one variable on the association of interest from the other (Mindt et al., 2008; Morton & Harper, 2007; Morton & Harper, 2009). According to

Morton & Harper (2007), "comparisons of bilingual and monolingual [individuals] drawn from immigrant and non-immigrant Canadian populations respectively are particularly hard to interpret because these populations differ in SES in complex but important ways. On the one hand, average family income is marginally lower for immigrant Canadian families than for non-immigrant Canadian families ($64,402 CAD versus $66,807 CAD, respectively according to the 2001 Canadian Census). On the other hand, immigrant Canadians are more educated than non-immigrant Canadians, due to an immigration policy that selects candidates on the basis of their academic achievement, language, and occupational skills" (p. 720). Based on this rationale, this study was also unique in its ability to control for SES and immigration independent of one another.

In spite of the potential lack of generalizability, there is no cause to believe that the AD etiological pathway, or multilingualism's relationship with this mechanism, would be altered in this study population. The results obtained by analysing this study population may facilitate the understanding of AD disease etiology, which can then later be assessed in more generalizable populations.

7.3.2 Access to Unique Covariate Data

Due to its thorough data collection process, the use of the Nun Study population allowed this investigation to examine the association between multilingualism and AD in the context of many novel covariates. The ability to evaluate the relationship between multilingualism and AD in the context of ApoE genotype was a large strength, given that E4 alleles are known to substantially increase AD risk. Moreover, our study found ApoE-E4 status to be one of few variables with statistically significant bearing on AD. This study is currently the only

investigation to consider the effects of ApoE-E4 status on the multilingualism-AD relationship, and these results suggest future studies would benefit from controlling for ApoE-E4 (and other genetic factors) when contemplating the role of multilingualism in AD development. Genetic influence on sporadic AD is not absolute, as many ApoE-E4 allele carriers do not develop AD and many AD cases do not have E4 alleles; however, given the relatively small number of well-established AD risk factors and the strength of this particular risk factor, it is important that the association between multilingualism and AD be evaluated in relation to genetics whenever possible.

The capacity to evaluate the association between multilingualism and AD in the context of linguistic ability was also a strength of this investigation. While multilingualism has been shown to confer certain cognitive advantages (e.g., in executive control), multilinguals have been shown to be disadvantaged in terms of verbal fluency (Craik & Bialystok, 2005; Rosselli et al., 2000). Multilinguals' vocabularies in each spoken language have been hypothesized to be less diverse compared to the vocabularies of monolinguals, as the time spent using each of the vocabularies is divided between the multiple spoken languages. According to Hakuta and Diaz (1985), multilinguals were thought to even have a "language handicap", exhibiting lower standards in writing composition and more grammatical errors than monolinguals (Hakuta & Diaz, 1985). Previous Nun Study investigations have demonstrated strong associations between high linguistic ability and a reduced AD risk (Riley et al., 2005; Snowdon et al., 1996; Snowdon et al., 2000). Therefore, linguistic ability has the potential to be a considerable confounding influence in any study of the relationship between multilingualism and AD. The Nun Study assessed linguistic ability using two different indicators: grammatical complexity, which was based on degree of sentence development, and idea density, which was measured using the

average number of ideas or emotions expressed per 10 words. These characteristics were assessed many years before AD development, as they were based on writing samples written during early adulthood. While neither idea density nor grammatical complexity measure verbal fluency exactly (phonetic or semantic category naming tests would provide more literal measures of verbal fluency), the current investigation is the first to evaluate the relationship between multilingualism and AD while controlling for any form of linguistic ability. Considering the intriguing results found by the linguistic ability sensitivity analyses, it is recommended that future prospective studies collect baseline verbal fluency information from participants (i.e., before clinical dementia manifests) so that it can be considered in assessments of multilingualism and AD.

As previously stated, the ability to assess the relationship between multilingualism and AD without the confounding influence of SES was a strength of this study. For instance, in 2007 Morton and Harper stressed that true cognitive effects of multilingualism may only be truly elucidated once investigators have controlled for SES differences. This was because much of the existing evidence in support of multilingual advantages had been derived from studies using multilingual and monolingual participants that may have differed considerably across SES levels and ethnicities. European studies of multilingual and monolinguals with small SES differences (compared to potential SES differences present in Canadian studies) have found it "notoriously difficult" to replicate the multilingual advantages in inhibitory control demonstrated by Canadian studies (Colzato et al., 2008). Our results might be explained in light of this discrepancy between study findings: a smaller observed effect of multilingualism may have been due to controlling for SES. Besides influencing education, nutrition, and mentally stimulating opportunities, SES has also been demonstrated to significantly affect an individual's

attentional control ability (Morton & Harper, 2007). These points are also especially relevant as all existing evidence (apart from the present study) supporting an association between multilingualism and a delay in dementia onset has been based on studies from Canada, particularly from large cities (Toronto (Bialystok et al., 2007; Craik et al., 2010) and Montreal (Chertkow et al., 2010)) where SES levels may not be uniform across all participants. Therefore, the SES comparability across all study participants was a considerable asset as it eliminated confounding from this source.

Another advantage of this study was that the participants had similar habits with respect to alcohol and tobacco use, making it possible to also control for these covariates. No other studies have accounted for these influences before. Given the effects these compounds can have on many other disease mechanisms and etiologies, the ability to investigate the association of interest without these external influences was advantageous.

The use of an outcome that considered AD neuropathology in addition to clinical dementia (with respect to Research Question 1) was also a novel study strength. Very few studies are able to gather post-mortem neuropathology information. Moreover, all measures of neuropathology were consistently assessed by a single neuropathologist who was blinded to participant cognitive status at the time of assessment. Therefore, it was not possible for neuropathology assessments to be biased by knowledge of subjects' clinical diagnoses. This also ensured that the application of the neuropathological criterion to each neural tissue sample was performed in a consistent fashion.

Without neuropathological confirmation, AD cannot be definitively identified. Many neurodegenerative disorders other than AD have the potential to present with clinical dementia, such as frontotemporal dementia and vascular dementia (Foster et al., 2007; Sultzer, Levin,

Mahler, High, & Cummings, 1993). Although less of a concern, other conditions such as multiple sclerosis may also present with clinical dementia similar to AD (Filley, Heaton, Nelson, Burks, & Franklin, 1989). Therefore, investigations that do not confirm clinical AD cases with ascertainment of AD neuropathology risk misclassification of study participants, limiting their potential to draw conclusions on AD-specific etiology.

7.3.3 Prospective Design of the Nun Study

The prospective nature of the original Nun Study bestowed many methodological strengths on the current investigation. Participants in the analytic sample were outcome-free at baseline; therefore, this investigation was able to ascertain incident cases. This is the first investigation to use prospective data in assessing multilingualism's association with AD and dementia. Therefore, it is also the first study in this research area to ascertain incident cases, which allowed for the direct calculation of risk. Furthermore, the ascertainment of incident cases also enabled this study to establish temporality between multilingualism and dementia development. The majority of the literature on multilingualism and AD is comprised of cross-sectional analyses which, despite being cost and time-effective, are limited by their ascertainment of prevalent dementia cases. The utilization of only prevalent cases is also problematic when the possibility of survival bias exists (as in the case of investigations concerning dementia and AD). As such, clinic-based studies may collect data only from cases surviving long enough to attend their memory clinic referral. Although the chances of this occurring often are low, survival bias would nevertheless be difficult to rule out in clinic-based studies. The present investigation was able to circumvent this problem by collecting data from every case that developed within the study population.

The ascertainment of incident cases by means of routine cognitive assessment also was advantageous to this study. Dementia diagnoses occurring in memory clinics may be differentially dependent on influences external to patient presentation, which subsequently may influence average age at diagnosis. For example, a dementia diagnosis often is made by a geriatric specialist, not a family physician. Therefore, the time at which a diagnosis is made depends on how long it takes for a patient to be seen by a physician (first by a family physician, and then by a referred specialist). In the province of Ontario, the wait times for geriatric referrals can be lengthy as there have been estimated to be only 300 geriatricians in Canada (as certified by the Royal College of Physicians and Surgeons of Canada in the specialty of Internal Medicine and the subspecialty of Geriatric Medicine; Mickleburgh, 2011). Therefore, the age at a dementia diagnosis may be artificially elevated for participants referred to specialists with longer waiting times. In our study, cognitive assessments occurred at relatively consistent intervals predetermined by study investigators; therefore, the time at which dementia diagnoses were made was independent of external influences present in clinic-based studies (e.g., referral time lengths). Moreover, our study was also able to identify cognitive impairment before it was recognized by even the participants themselves, which led to reduced delays between initial problem development and decisions on whether to seek care. Therefore, this study had the capacity to more accurately compare participants' estimated time of dementia development compared to other studies of this event, since cognitive assessments were performed during every wave of follow-up. Additionally, cognitive assessments were performed on each study participant during every follow-up visit, regardless of their prior cognitive state. This allowed for a wealth of information to be collected about participants without dementia, which also made our analyses more meaningful as they factored in these data as well as data from dementia

cases into calculations of dementia likelihood. Additionally, since our study utilized data collected from cognitive assessment information in approximately consistent intervals (mean length of time between consecutive cognitive assessments = 1.54 years; SD = 0.32 years), analyses using discrete-time survival analysis models were valid.

The systematic compilation of data on every study participant, regardless of their cognitive status, ensured that our study was less susceptible to information bias. For instance, if an investigator is aware of a participant's outcome status they may be more likely to probe for more detailed information concerning variables of interest. In this study, however, data concerning covariates of interest were collected at baseline (i.e., when all participants were dementia and AD-free). Moreover, data on multilingualism were ascertained several years prior to study inception, which eliminated the potential for prevarication bias (the selective revealing or suppression of information). Cross-sectional studies are particularly susceptible to this kind of information bias, as participants may be familiar with a given investigator's past research or current research aim. Past cross-sectional studies have also collected data from participants already having dementia and thus the accuracy of these self-report data may be questionable. In order to avoid this problem, these past reports have also utilized proxy reports as means of data collection, which may result in exposure misclassification (Nelson, Longstreth, Koepsell, Checkoway, & van Belle, 1994).

The use of data from all possible participants also ensured this study would be less vulnerable to selection bias. Studies using data collected from only patients presenting at memory clinics are at risk for selection bias, since patients with more severe forms of clinical dementia or other comorbidities may not be able to participate in studies, meaning such investigations might not be able to capture information concerning some of the most afflicted

individuals. Furthermore, it is conceivable that many dementia cases may not present to AD clinics (e.g., limited access to medical attention; death before they are able to be scheduled for a referral to a specialist; non-compliance), which means that data on these types of individuals could not be considered by clinic-based studies.

Investigations employing data collected from only AD cases are unable to compare these participants to individuals who do not develop AD (or dementia). In addition to the lack of a reference group, it is conceivable that individuals not developing a given disease may provide great insight concerning disease avoidance and protective factors. Therefore, studies that do not incorporate individuals who successfully avoid AD or dementia may risk the omission of data concerning protective (or risk) factors. This illustrates the advantage available to our investigation as it could analyse data collected from participants who developed dementia as well as from those who did not.

Lastly, since data were available on the participants excluded from the current investigation's analytic samples, the characteristics of participants included in the analyses were able to be compared with those of the excluded participants. Investigations not employing a prospective cohort design, conversely, would be limited in their ability to perform this evaluation of non-response, as data concerning excluded individuals or drop-outs would be unavailable. Therefore, while this study observed some significant differences between the characteristics of included and excluded participants, the ability to detect and acknowledge these differences allowed for our results to be contextually interpreted.

7.3.4 Advantages of Secondary Data

As mentioned previously, the analysis of secondary data is often accompanied by various methodological limitations. In the instance of this investigation, however, the methodological

advantages gained by employing secondary data from the Nun Study outweigh the limitations. Firstly, the execution of a prospective cohort study requires an extensive amount of time, money, and resources. This study collected 12 waves of cognitive and physical assessment data, in addition to participant blood samples and genetic information. Finally, after participants were followed until death, this study was permitted to examine participant brains and evaluate neurological data (which included AD neuropathology in addition to many other neurological features critical to the research of other disease mechanisms). Furthermore, since the Nun Study also had access to convent archives, this investigation had the ability to use historical data on study participants, collected decades before the conception of the study questions. Another strength of using these secondary data was that all study measures were developed with the deliberate aim of researching aging and AD. While the association between multilingualism and AD or dementia was not among the study questions originally conceptualized by investigators, the original primary Nun Study investigators ensured data on many factors relevant to AD and dementia were collected. This foresight allowed our study the advantage of applying several new covariates, such as ApoE-E4 allele status and linguistic ability, to this particular research question.

7.4 Implications and Future Research Directions

Currently 1 in 11 Canadians over age 65 have dementia (Smetanin et al., 2009) and more than 35 million people have dementia worldwide (Wimo & Prince, 2010). Approximately $604 billion dollars US were spent in 2010 on direct care, caregiving, and home care for dementia patients across the globe (Wimo & Prince, 2010); therefore, as the number and proportion of individuals aged 65 and older steadily grows, AD and dementia will increasingly burden our health care system. In addition to these financial costs, AD and dementia represent sources of overwhelming emotional burden and distress. As the number of individuals aged 65 and older steadily grows, AD and dementia will continue to burden our health care system and be a source of emotional stress. Given that AD and dementia currently have no known cure, the best approach to addressing this problem is by means of prevention. If dementia and AD are to be prevented, risk factors must first be identified so they potentially may be avoided. Similarly, if protective factors against dementia and AD can be determined, individuals may be able to reduce their risk despite pre-existing biological predispositions. If new insights on the relationship between multilingualism and late-life cognitive outcomes can be uncovered, a strategy could be developed for AD or dementia prevention by means of gaining additional language proficiency. Similarly, elucidating the association between multilingualism and cognitive reserve will improve our understanding of cognitive reserve and its contributing factors, which ultimately may lead to an overall avoidance strategy for cognitive decline.

The first aim of this investigation was to evaluate the association between multilingualism and AD. Before controlling for linguistic ability indicators, multilingualism did not appear to have any significant association with AD. However, multilingualism did appear to be associated with a decreased AD likelihood once the analyses adjusted for grammatical complexity.

Therefore, the collection of information on grammatical complexity is important to the investigation of multilingualism's association with AD risk. The results of our analyses also reiterate the importance of the ApoE-E4 allele as an AD risk factor. Based on the established body of evidence concerning ApoE-E4 and AD risk, in the future individuals may wish to know whether they possess an ApoE-E4 allele so that they might become more knowledgeable of their personal AD risk profiles. Furthermore, with this information one may become motivated to pursue certain mental engagements in order to decrease AD susceptibility. The second aim of this study was to evaluate whether proficiencies in differing numbers of languages affected the time at which one became more likely to develop clinical dementia. Participants with differing language proficiencies were found to have different estimates of dementia likelihood; however, many of the associations contradicted the *a priori* hypothesis that monolinguals would have the highest susceptibility to dementia. This study did, however, confirm the protective role of education against dementia development. This suggests that, while one pursues education for many reasons other than dementia prevention, one may consider trying to improve their chances of living dementia-free in late life by pursuing educational activities.

The most notable study finding was that linguistic ability (grammatical complexity and idea density) appeared to change both the association between multilingualism and AD and the association between multilingualism and dementia. When evaluated in the context of grammatical complexity, multilingualism displayed an association with reduced AD likelihood; similarly, when multilingualism was evaluated in the context of idea density, proficiency in three or more languages was associated with reduced dementia likelihood. These findings suggest that, when linguistic ability differences between participants are not considered,

inherent multilingual disadvantages in linguistic ability may conceal the true protective association of multilingualism with AD or dementia. This may occur due to the strong relationship between poor linguistic ability and increased AD likelihood, as demonstrated by other Nun Study investigations (Riley et al., 2005; Snowdon et al., 1996).

It is fair to propose that multilingualism may enhance cognitive reserve due to multilingual advantages in executive control (Bialystok et al., 2007). However, one must also bear in mind the suggested accompanying cognitive disadvantages of multilingualism when evaluating the relationship between multilingualism and late-life cognitive outcomes. Therefore, in order for the association between multilingualism and dementia to be accurately measured in the future, data concerning suggested multilingual disadvantages (i.e., deficits in verbal fluency (Bialystok et al., 2008; Bialystok, 2009; Gollan, Montoya, Cera, & Sandoval, 2008; Rosselli et al., 2000); less vocabulary development (Bialystok, 2009; Mahon & Crutchley, 2006; Oller & Eilers, 2002)) should also be considered.

While some of our findings support an association between multilingualism and protection against dementia and AD, this protective effect was not consistently observed throughout the entire study. In fact, speaking two languages was repeatedly associated with a non-significant increased likelihood for both AD and dementia, which conflicts with previous reports of protective associations between bilingualism and dementia (Bialystok et al., 2007; Craik et al., 2010). This discrepancy between findings may have resulted for a variety of reasons. Firstly, our study analyzed data on both individuals who had dementia as well as those who did not develop dementia when evaluating the relationship between multilingualism and dementia or AD. Since the AD or dementia-associated benefits of multilingualism have been studied in only participants with dementia, it is understandable that different findings were

derived when analyses also considered individuals remaining dementia-free. Secondly, our study was able to consider many potential confounding influences when evaluating the associations of both dementia and AD with multilingualism. Past findings have not accounted for ApoE-E4 status, nor have they evaluated multilingualism's relationship with dementia in the context of grammatical complexity and idea density. All three of these variables were found by our analyses to be significantly related to the estimation of AD, and dementia, likelihood. Therefore, it is understandable that different findings resulted from unadjusted calculations. Lastly, while past findings have reported a protection from multilingualism against AD (Chertkow et al., 2010; Craik et al., 2010), these investigators did not consider neuropathology in their outcome definitions and thus were only truly evaluating the relationship between multilingualism and clinical dementia. Our findings showed similarities to those reported previously in the literature, given our evaluation of the relationship of multilingualism with AD and dementia suggested that speaking four or more languages was associated with decreased dementia risk.

Convening lines of preliminary evidence on enhanced reserve and high cognitive stimulation support further epidemiological exploration and hypothesis generation in this research area. Multilingualism is a plausible example of mentally stimulating activity; nevertheless, the association between multilingualism and late-life cognitive decline still needs further elucidation. Future investigations in this research area should be methodologically diverse so new insights may be gained and theories can be formulated. Data regarding mental stimulation or cognitive reserve to date remain insufficient to postulate any practical public health initiatives; however, it still represents a promising avenue for protective factor research, as even non-significant findings may help to clarify new relationships. New findings relating to

cognitive reserve may also help to clarify existing relationships between mental stimulation and late-life cognition. The findings of this investigation may help to direct future endeavours, if they can be replicated and confirmed in other populations using a similar variety of covariates.

References

Akiyama, H., Barger, S., Barnum, S., Bradt, B., Bauer, J., Cole, G. M., et al. (2000). Inflammation and Alzheimer's disease. *Neurobiology of Aging, 21*(3), 383-421.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders, (4th edition).* Washington, DC: American Psychiatric Association.

Baas, P. W., & Qiang, L. (2005). Neuronal microtubules: When the MAP is the roadblock. *Trends in Cell Biology, 15*(4), 183-187.

Bartus, R., Dean, R., Beer, B., & Lippa, A. (1982). The cholinergic hypothesis of geriatric memory dysfunction. *Science, 217*(4558), 408-414.

Bartzokis, G. (2004). Age-related myelin breakdown: A developmental model of cognitive decline and Alzheimer's disease. *Neurobiology of Aging, 25*(1), 5-18.

Bartzokis, G., Cummings, J. L., Sultzer, D., Henderson, V. W., Nuechterlein, K. H., & Mintz, J. (2003). White matter structural integrity in healthy aging adults and patients with Alzheimer disease: A magnetic resonance imaging study. *Archives of Neurology, 60*(3), 393-398.

Bartzokis, G., Sultzer, D., Lu, P. H., Nuechterlein, K. H., Mintz, J., & Cummings, J. L. (2004). Heterogeneous age-related breakdown of white matter structural integrity: Implications for cortical "disconnection" in aging and Alzheimer's disease. *Neurobiology of Aging, 25*(7), 843-851.

Bermejo-Pareja, F., Benito-Leon, J., Vega, S., Medrano, M. J., & Roman, G. C. (2008). Incidence and subtypes of dementia in three elderly populations of central Spain. *Journal of the Neurological Sciences, 264*(1-2), 63-72.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*(1), 3-11.

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

Bialystok, E., Craik, F., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging, 19*(2), 290-303.

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition, 34*(4), 859-873.

Bialystok, E., Craik, F. I., & Ryan, J. (2006). Executive control in a modified anti-saccade task: Effects on aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory and Cognition, 32*, 1341-1354.

Bialystok, E., & DePape, A. (2009). Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology, 35*(2), 565-574.

Bialystok, E., & Majumder, S. (1998). The relationship between bilingualism and the development of cognitive processes in problem solving. *Applied Psycholinguistics, 19*, 69-85.

Bird, T. D., Stranahan, S., Sumi, S. M., & Raskind, M. (1983). Alzheimer's disease: Choline acetyltransferase activity in brain tissue from clinical and pathological subgroups. *Annals of Neurology, 14*(3), 284-293.

Blennow, K., de Leon, M. J., & Zetterberg, H. (2006). Alzheimer's disease. *The Lancet, 368*, 387-403.

Bochner, S. (1996). The learning strategies of bilingual versus monolingual students. *British Journal of Educational Psychology, 66*, 83-93.

Borchelt, D. R., Ratovitski, T., van Lare, V., Lee, M. K., Gonzales, V., Jenkins, N. A., et al. (1997). Accelerated amyloid deposition in the brains of transgenic mice coexpressing mutant presenilin 1 and amyloid precursor proteins. *Neuron, 19*(4), 939-945.

Borenstein, A. R., Wu, Y., Mortimer, J. A., Schellenberg, G. D., McCormick, W. C., Bowen, J. D., et al. (2005). Developmental and vascular risk factors for Alzheimer's disease. *Neurobiology of Aging, 26*(3), 325-34.

Bowen, D. M., Allen, S. J., Benton, J. S., Goodhardt, M. J., Haan, E. A., Palmer, A. M., et al. (1983). Biochemical assessment of serotonergic and cholinergic dysfunction and cerebral atrophy in Alzheimer's disease. *Journal of Neurochemistry, 41*(1), 266-272.

Braak, E., Griffing, K., Arai, K., Bohl, J., Bratzke, H., & Braak, H. (1999). Neuropathology of Alzheimer’s disease: What is new since A. Alzheimer? *European Archives of Psychiatry and Clinical Neuroscience, 249*(S3), 14-22.

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

Braak, H., Del Tredici, K., Rüba, U., de Vos, R. A. I., Steur, E. N. H., & Braak, E. (2003). Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiology of Aging, 24*, 197-211.

Breteler, M. M. B. (2000). Vascular risk factors for Alzheimer's disease: An epidemiologic perspective. *Neurobiology of Aging, 21*, 153-160.

Burke, D. M., & Shafto, M. A. (2008). Language and aging. In F. Craik & T. Salthouse (Eds.), *The handbook of aging and cognition* (3rd ed., pp. 373-443). New York, NY: Psychology Press.

Busciglio, J., Lorenzo, A., Yeh, J., & Yankner, B. A. (1995). β-Amyloid fibrils induce tau phosphorylation and loss of microtubule binding. *Neuron, 14*(4), 879.

Butterfield, D. A., & Lauderback, C. M. (2002). Lipid peroxidation and protein oxidation in Alzheimer's disease brain: Potential causes and consequences involving amyloid β-peptide-associated free radical oxidative stress. *Free Radical Biology and Medicine, 32*(11), 1050-1060.

Carr, D. B., Goate, A., Phil, D., & Morris, J. C. (1997). Current concepts in the pathogenesis of Alzheimer's disease. *The American Journal of Medicine, 103*(3, Supplement 1), 3S-10S.

Cataldo, J. K., Prochaska, J. J., & Glantz, S. A. (2010). Cigarette smoking is a risk factor for Alzheimer's disease: An analysis controlling for tobacco industry affiliation. *Journal of Alzheimer's Disease, 19*(2), 465.

Chertkow, H., Whitehead, V., Wolfson, C., Atherton, J., & Bergman, H. (2010). Multilingualism (but not always bilingualism) delays the onset of Alzheimer disease: Evidence from a bilingual community. *Alzheimer Disease and Associated Disorders, 24*(2), 118-125.

Cheung, H., & Kemper, S. (1992). Competing complexity metrics and adults' production of complex sentences. *Applied Psycholinguistics, 13*(1), 53-76.

Citron, M., Oltersdorf, T., & Haass, C. (1992). Mutation of the β-amyloid precursor protein in familial Alzheimer's disease increases β-protein production. *Nature, 360,* 672-674.

Colzato, L. S., Bajo, M.T. van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W., & Hommel, B. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(2), 302-312.

Corder, E. H., Saunders, A. M., & Strittmater, W. J. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science, 261,* 921-923.

Coyle, J., Price, D., & DeLong, M. (1983). Alzheimer's disease: A disorder of cortical cholinergic innervation. *Science, 219*(4589), 1184-1190.

Craik, F., & Bialystok, E. (2005). Intelligence and executive control: Evidence from aging and bilingualism. *Cortex, 41*, 222-224.

Craik, F., & Bialystok, E. (2006). Planning and task management in older adults: Cooking breakfast. *Memory and Cognition, 34*(6), 1236-1249.

Craik, F. I. M., Bialystok, E., & Freedman, M. (2010). Delaying the onset of Alzheimer disease. *Neurology, 75*(19), 1726-1729.

Crystal, H., Dickson, D., Fuld, P., Masur, D., Scott, R., Mehler, M., et al. (1988). Clinico-pathologic studies in dementia: Nondemented subjects with pathologically confirmed Alzheimer's disease. *Neurology, 38*(11), 1682.

Cummings, J. L., & Kaufer, D. (1996). Neuropsychiatric aspects of Alzheimer's disease: The cholinergic hypothesis revisited. *Neurology, 47*(4), 876-883.

Daffner, K. R. (2010). Promoting successful cognitive aging: A comprehensive review. *Journal of Alzheimer's Disease, 19*, 1101-1122.

de la Monte, S. M. (1989). Quantitation of cerebral atrophy in preclinical and end-stage Alzheimer's disease. *Annals of Neurology, 25*(5), 450-459.

de la Torre, J. C. (2002). Alzheimer disease as a vascular disorder: Nosological evidence. *Stroke, 33*(4), 1152-1162.

de la Torre, J. C. (2004). Is Alzheimer's disease a neurodegenerative or a vascular disorder? Data, dogma, and dialectics. *The Lancet Neurology, 3*(3), 184-190.

de la Torre, J. C. (2010). Vascular risk factor detection and control may prevent Alzheimer's disease. *Ageing Research Reviews, 9*(3), 218-225.

DeKosky, S. T., Ikonomovic, M. D., Styren, S. T., Beckett, L., Wisniewski, S., Bennett, D. A., et al. (2002). Upregulation of choline acetyltransferase activity in hippocampus and frontal cortex of elderly subjects with mild cognitive impairment. *Annals of Neurology, 51*(2), 145-155.

Drachman, D. A., & Leavitt, J. (1974). Human memory and the cholinergic system: A relationship to aging? *Archives of Neurology, 30*(2), 113-121.

Dubois, B., Feldman, H. H., Jacova, C., Dekosky, S. T., Barberger-Gateau, P., Cummings, J., et al. (2007). Research criteria for the diagnosis of Alzheimer's disease: Revising the NINCDS-ADRDA criteria. *Lancet Neurology, 6*(8), 734-746.

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B. & Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science, 270*, 305–307.

Esch, F. S., Keim, P. S., Beattie, E. C., Blacher, R. W., Culwell, A. R., Oltersdorf, T., et al. (1990). Cleavage of amyloid β peptide during constitutive processing of its precursor. *Science, 248*(4959), 1122-1124.

Estus, S., Golde, T. E., Kunishita, T., Blades, D., Lowery, D., Eisen, M., et al. (1992). Potentially amyloidogenic, carboxyl-terminal derivatives of the amyloid protein precursor. *Science, 755*, 726-728.

Evans, D. A., Beckett, L. A., Albert, M. S., Hebert, L. E., Scherr, P. A., Funkenstein, H. H., et al. (1993). Level of education and change in cognitive function in a community population of older persons. *Annals of Epidemiology, 3*(1), 71-7.

Evans, M. A., Kim, H. M., & O'Brien, T. E. (1996). An application of profile-likelihood confidence interval to capture-recapture estimators. *Journal of Agricultural, Biological and Environmental Statistics, 1*(1), 131-140.

Farlow, M. R. (2010). Alzheimer's disease. In H. M. Fillit, K. R. Rockwood & K. Woodhouse (Eds.), *Brocklehurst's textbook of geriatric medicine and gerontology* (7th ed., pp. 411). Philadelphia, PA: Saunders Elsevier.

Farrer, L. A., Cupples, L. A., Haines, J. L., Hyman, B., Kukull, W. A., Mayeux, R., et al. (1997). Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease: A meta-analysis. *JAMA: The Journal of the American Medical Association, 278*(16), 1349-1356.

Fernandes, M. A., Craik, F., Bialystok, E., & Kreuger, S. (2007). Effects of bilingualism, aging, and semantic relatedness on memory under divided attention. *Canadian Journal of Experimental Psychology, 61*(2), 128-141.

Ferri, C. P., Prince, M., Brayne, C., Brodaty, H., Fratiglioni, L., Ganguli, M., et al. (2005). Global prevalence of dementia: A delphi consensus study. *The Lancet, 366*(9503), 2112-2117.

Filley, C. M., Heaton, R. K., Nelson, L. M., Burks, J. S., & Franklin, G. M. (1989). A comparison of dementia in Alzheimer's disease and multiple sclerosis. *Archives of Neurology, 46*(2), 157-161.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189-198.

Foot, D. K. (2008). *Some economic and social consequences of population aging.* No. 7. Montreal, QC: Institute for Research on Public Policy. Retrieved from: http://www.irpp.org/cpa/briefs/foot.pdf

Foster, N. L., Heidebrink, J. L., Clark, C. M., Jagust, W. J., Arnold, S. E., Barbas, N. R., et al. (2007). FDG-PET improves accuracy in distinguishing frontotemporal dementia and Alzheimer's disease. *Brain, 130*(10), 2616-2635.

Francis, P. T., Palmer, A. M., Snape, M., & Wilcock, G. K. (1999). The cholinergic hypothesis of Alzheimer's disease: A review of progress. *Journal of Neurology, Neurosurgery & Psychiatry, 66*(2), 137-147.

Fratiglioni, L., Grut, M., Forsell, Y., Viitanen, M., Grafström, M., Holmen, K., et al. (1991). Prevalence of Alzheimer's disease and other dementias in an elderly urban population. *Neurology, 41*(12), 1886-1886.

Fratiglioni, L., Paillard-Borg, S., & Winblad, B. (2004). An active and socially integrated lifestyle in late life might protect against dementia. *Lancet Neurology, 3*, 343-353.

Fratiglioni, L., Viitanen, M., von Strauss, E., Tontodonati, V., Herlitz, A., & Winblad, B. (1997). Very old women at highest risk of dementia and Alzheimer's disease: Incidence data from the Kungshomen project, Stockholm. *Neurology, 48*, 132-138.

Friis, H., & Sellers, T. (2009). *Epidemiology for Public Health Practice* (4th ed.). Sudbury, MA: Jones and Bartlett Publishers.

Gao, S., Hendrie, H. C., Hall, K. S., & Hui, S. (1998). The relationships between age, sex, and the incidence of dementia and Alzheimer disease. *Archives of General Psychiatry, 55*, 809-815.

Gatz, M., Svedberg, P., Pedersen, N. L., Mortimer, J. A., Berg, S., & Johansson, B. (2001). Education and the risk of Alzheimer's disease: Findings from the study of dementia in Swedish twins. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 56B*(5), 292-300.

Giannakopoulos, P., Herrmann, F., Bussière, T., Bouras, C., Kövari, E., Perl, D., et al. (2003). Tangle and neuron numbers, but not amyloid load, predict cognitive status in Alzheimer's disease. *Neurology, 60*(9),1495-1500.

Golde, T. E., Estus, S., & Younkin, L. H. (1992). Processing of the amyloid protein precursor to potentially amyloidogenic derivatives. *Science, 255*, 728-730.

Gollan, T., Montoya, R., Cera, C., & Sandoval, T. (2008). More use always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language, 58*, 787-814.

Grundke-Iqbal, I., Iqbal, K., Tung, Y. C., Quinlan, M., Wisniewski, H. M., & Binder, L. I. (1986). Abnormal phosphorylation of the microtubule-associated protein tau (tau) in

Alzheimer cytoskeletal pathology. *Proceedings of the National Academy of Sciences of the United States of America, 83*(13), 4913-4917.

Haass, C., Schlossmacher, M. G., Hung, A. Y., Vig-Pelfrey, C., Mellon, A., Ostaszewski, B. L., et al. (1992). Amyloid beta -peptide is produced by cultured cells during normal metabolism. *Nature, 359*(6393), 322-325.

Hakuta, K., & Diaz, R. M. (1985). The relationship between degree of bilingualism and cognitive ability: A critical discussion and some new longitudinal data. In K. E. Nelson, & A. van Kleeck (Eds.), *Children's language* (pp. 320). Hillsdale, NJ: Lawrence Erlbraum and Associates.

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science, 14*(1), 31-38.

Hardy, J. (1997). Amyloid, the presenilins and Alzheimer's disease. *Trends in Neurosciences, 20*(4), 154-159.

Hardy, J. (2009). The amyloid hypothesis for Alzheimer's disease: A critical reappraisal. *Journal of Neurochemistry, 110*(4), 1129-1134.

Hardy, J. A., & Higgins, G. A. (1992). Alzheimer's disease: The amyloid cascade hypothesis. *Science, 256*(5054), 184-185.

Hardy, J., & Selkoe, D. J. (2002). The amyloid hypothesis of Alzheimer's disease: Progress and problems on the road to therapeutics. *Science, 297*(5580), 353.

Harvey, R. J., Skelton-Robinson, M., & Rossor, M. N. (2003). The prevalence and causes of dementia in people under the age of 65 years. *Journal of Neurology, Neurosurgery & Psychiatry, 74*(9), 1206-1209.

Health Canada. (2001). *Health care expenditures in Canada by age and sex, 1980 to 2000-01.*(H21-172/2001). Ottawa, ON: Health Canada. Retrieved from http://www.hc-sc.gc.ca/hcs-sss/pubs/expen-depens/index-eng.php.

Holmes, C., Boche, D., Wilkinson, D., Yadegarfar, G., Hopkins, V., Bayer, A., et al. (2008). Long-term effects of Aβ42 immunisation in Alzheimer's disease: Follow-up of a randomised, placebo-controlled phase I trial. *The Lancet, 372*(9634), 216-223.

Holtzman, D. M., Fagan, A. M., Mackey, B., Tenkova, T., Sartorius, L., Paul, S. M., et al. (2000). Apolipoprotein E facilitates neuritic and cerebrovascular plaque formation in an Alzheimer□ s disease model.*Annals of Neurology, 47*(6), 739-747.

Iqbal, K., del C. Alonso, A., Chen, S., Chohan, M. O., El-Akkad, E., Gong, C., et al. (2005). Tau pathology in Alzheimer disease and other tauopathies. *Biochimica Et Biophysica Acta (BBA) - Molecular Basis of Disease, 1739*(2-3), 198-210.

Irizarry, M. C., Soriano, F., McNamara, M., Page, K. J., Schenk, D., Games, D., et al. (1997). A-beta deposition is associated with neuropil changes, but not with overt neuronal loss in the human amyloid precursor protein V717F (PDAPP) transgenic mouse. *Journal of Neuroscience, 17*(18), 7053-7059.

Janicki, S. C., & Schupf, N. (2010). Hormonal influences on cognition and risk for Alzheimer's disease. *Current Neurology and Neuroscience Reports, 10*(5), 359-366.

Jeste, D. V., Meeks, T. W., Kim, D. S., & Zubenko, G. S. (2006). Research agenda for DSM-V: Diagnostic categories and criteria for neuropsychiatric syndromes in dementia. *Journal of Geriatric Psychiatry and Neurology, 19*(3), 160-171.

Jorm, A. F., Korten, A. E., & Henderson, A. S. (1987). The prevalence of dementia: A quantitative integration of the literature. *Acta Psychiatrica Scandinavica, 76*(5), 465-479.

Kalaria, R. N. (2003). Vascular factors in Alzheimer's disease. *International Psychogeriatrics, 15*(S1), 47-52.

Kang, J., Lemaire, H., Unterbeck, A., Salbaum, J. M., Masters, C. L., Grzeschik, K., et al. (1987). The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor. *Nature, 325*, 733-736.

Karp, A., Kåreholt, I., Qiu, C., Bellander, T., Winblad, B., & Fratiglioni, L. (2004). Relation of education and occupation-based socioeconomic status to incident Alzheimer's disease. *American Journal of Epidemiology, 160*(4), 404-5.

Katzman, R., Terry, R., DeTeresa, R., Brown, T., Davies, P., Fuld, P., et al. (1988). Clinical, pathological, and neurochemical changes in dementia: A subgroup with preserved mental status and numerous neocortical plaques. *Annals of Neurology, 23*(2), 138-144.

Katzman, R., Brown, T., Fuld, P., Peck, A., Schechter, R., & Schimmel, H. (1983). Validation of a short orientation-memory-concentration test of cognitive impairment. *American Journal of Psychiatry, 140*(6), 734-739.

Kavé, G., Eyal, N., Shorek, A., & Cohen-Mansfield, J. Multilingualism and cognitive states in the oldest old. *Psychology and Aging, 23*(1), 70-78.

Kennedy, W. J., & Bancroft, T. A. (1971). Model building for prediction in regression based upon repeated significant tests. *Annals of Mathematical Statistics, 42*(4), 1273-1284.

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.

Knopman, D. S., Parisi, J. E., Salviati, A., Floriach-Robert, M., Boeve, B. F., Ivnik, R. J., et al. (2003). Neuropathology of cognitively normal elderly. *Journal of Neuropathology and Experimental Neurology, 62*(11), 1087-1095.

Kolata, G. (2010, August 4). In push to detect early Alzheimer's markers, hopes for prevention. *New York Times.* Retrieved May 1, 2011, from http://www.nytimes.com/2010/08/05/health/05alzheimers.html

Kukull, W. A., Larson, E. B., Teri, L., Bowen, J., McCormick, W., & Pfanschmidt, M. L. (1994). The Mini-Mental State Examination score and the clinical diagnosis of dementia. *Journal of Clinical Epidemiology, 47*(9), 1061-1067.

Kuriansky, J., & Gurland, B. (1976). The performance test of activities of daily living. *International Journal of Aging and Human Development, 73*(4), 343-352.

Kurt, M. A., Davies, D. C., Kidd, M., Duff, K., Rolph, S. C., Jennings, K. H., et al. (2001). Neurodegenerative changes associated with β-amyloid deposition in the brains of mice carrying mutant amyloid precursor protein and mutant presenilin-1 transgenes. *Experimental Neurology, 171*(1), 59-71.

Launer, L. J., Andersen, K., Dewey, M. E., Letenneur, L., Ott, A., Amaducci, L. A., et al. (1999). Rates and risk factors for dementia and Alzheimer's disease: Results from EURODEM pooled analyses. *Neurology, 52*(1), 78-84.

Le Carret, N., Lafont, S., Letenneur, L., Dartigues, J., Mayo, W., & Fabrigoule, C. (2003). The effect of education on cognitive performances and its implication for the constitution of the cognitive reserve. *Developmental Neuropsychology, 23*(3), 317-337.

Lee, J. H. (2003). Genetic evidence for cognitive reserve: Variations in memory and related cognitive functions. *Journal of Clinical and Experimental Neuropsychology, 25*(5), 594-613.

Lee, K. I., & Koval, J. J. (1997). Determination of the best significance level in forward stepwise logistic regression. *Communications in Statistics - Simulation and Computation, 26*(2), 559-575.

Lindsay, J., Laurin, D., Verreault, R., Hébert, R., Helliwell, B., Hill, G. B., et al. (2002). Risk factors for Alzheimer's disease: A prospective analysis from the Canadian Study of Health and Aging. *American Journal of Epidemiology, 156*(5), 445-453.

Maguire, E., Gadian, D., Johnsrude, I., Good, C., Ashburner, J., Frackowiak, R., & Frith, C. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences of the United States of America, 97*, 4398–4403.

Mahon, M., & Crutchley, A. (2006). Performance of typically-developing school-age children with English as an additional language on the British picture vocabulary scales II. *Child Language Teaching and Therapy, 22*(3), 333-351.

Manly, J. J., Schupf, N., Tang, M. X., & Stern, Y. (2005). Cognitive decline and literacy among ethnically diverse elders. *Journal of Geriatric Psychiatry and Neurology, 18*, 213-217.

Markesbery, W. R. (1997). Oxidative stress hypothesis in Alzheimer's disease. *Free Radical Biology and Medicine, 23*(1), 134-147.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. (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*, 939-44.

Mechelli, A., Crinion, J. T., Noppeny, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., et al. (2004). Neurolinguistics: Structural plasticity in the bilingual brain. *Nature, 431*, 757.

Mickleburgh, R. (2011, March 24). Are we in denial about the growing issue of elder care? *The Globe and Mail*. Retrieved May 7, 2011, from http://www.theglobeandmail.com/life/health/new-health/health-policy/are-we-in-denial-about-the-growing-issue-of-elder care/article1955902/

Mindt, M. R., Arentoft, A., Germano, K. K., D'Aquila, E., Scheiner, D., Pizzirusso, M., et al. (2008). Neuropsychological, cognitive, and theoretical considerations for evaluation of bilingual individuals. *Neuropsychology Review, 18*, 255-268.

Mirra, S. S., Heyman, A., McKeel, D., Sumi, S. M., Crain, B. J., Brownlee, L. M., et al. (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.

Moceri, V. M., Kukull, W. A., Emanuel, I., van Belle, G., & Larson, E. B. (2000). Early-life risk factors and the development of Alzheimer's disease. *Neurology, 54*(2), 415-20.

Moceri, V. M., Kukull, W. A., Emanual, I., van Belle, G., Starr, J. R., Schellenberg, G. D., et al. (2001). Using census data and birth certificates to reconstruct the early-life socioeconomic environment and the relation to the development of Alzheimer's disease. *Epidemiology, 12*(4), 383-9.

Morris, J. C., Heyman, A., Mohs, R. C., & Hughes, J. P. (1989). The Consortium to Establish a Registry for Alzheimer's disease (CERAD): I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology, 39*(9), 1159-1165.

Mortimer, J. A., Snowdon, D. A., & Markesbery, W. R. (2003). Head circumference, education and risk of dementia: Findings from the Nun Study. *Journal of Clinical and Experimental Neuropsychology, 25*(5), 671.

Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science, 10*(6), 719-726.

Morton, J. B., & Harper, S. N. (2009). Bilinguals show an advantage in cognitive control - the question is why. *Developmental Science, 12*(4), 502-503.

Muthén, B., & Masyn, K. (2005). Discrete-time survival mixture analysis. *Journal of Educational and Behavioral Statistics, 30*(1), 27-58.

Nagai, T., McGeer, P. L., Peng, J. H., McGeer, E. G., & Dolman, C. E. (1983). Choline acetyltransferase immunohistochemistry in brains of Alzheimer's disease patients and controls. *Neuroscience Letters, 36*(2), 195-199.

The National Institute on Aging - Reagan Institute (NIA-RI) Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease. (1997). Consensus recommendations for the postmortem diagnosis of Alzheimer's disease. *Neurobiology of Aging, 18*(4 Suppl), S1-2.

The National Institute on Aging / Alzheimer's Association Workgroup. (2010). New criteria and guidelines for the diagnosis of Alzheimer's disease published for the first time in 27 years. *News release from the Alzheimer's Association.* Chicago, IL: The Alzheimer's Association. Retrieved from: http://wwwalz.org/documents_custom/Alz_Assoc_ diag_criteria_guidelines_press_release_041911.pdf

Nelson, P. T., Kukull, W. A., & Frosch, M. P. (2010). Thinking outside the box: Alzheimer-type neuropathology that does not map directly onto current consensus recommendations. *Journal of Neuropathology and Experimental Neurology, 69*(5), 449-454.

Nelson, L. M., Longstreth, W. T., Koepsell, T. D., Checkoway, H., & van Belle, G. (1994). Completeness and accuracy of interview data from proxy respondents: Demographic, medical, and life-style factors. *Epidemiology*, *5*(2), 204-217.

Nilsson, L., Nordberg, A., Hardy, J., Wester, P., & Winblad, B. (1986). Physostigmine restores ^{3}H-acetylcholine efflux from Alzheimer brain slices to normal level. *Journal of Neural Transmission, 67*, 275-285.

Oller, D. K., & Eilers, R. E. (2002). *Language and literacy in bilingual children*. Clevedon, UK: Multilingual Matters.

Panza, F., Capurso, C., D'Introno, A., Colacicco, A. M., Frisardi, V., Santamato, A., et al. (2008). Vascular risk factors, alcohol intake, and cognitive decline. *The Journal of Nutrition, Health and Aging, 12*(6), 376.

Paradis, M. (2005). Aspects and implications of bilingualism. In J. Kroll, & A. M. B. DeGroot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 411). New York, NY: Oxford University Press.

Perry, E. K., Gibson, P. H., Blessed, G., Perry, R. H., & Tomlinson, B. E. (1977). Neurotransmitter enzyme abnormalities in senile dementia: Choline acetyltransferase and glutamic acid decarboxylase activities in necropsy brain tissue. *Journal of the Neurological Sciences, 34*(2), 247-265.

Potvin, A. R., Tourtellotte, W. W., Dailey, J. S., Alberts, J. W., Walker, J. E., Pew, R. W., et al. (1972). Simulated activities of daily living examination. *Archives of Physical Medicine and Rehabilitation, 53*(10), 476-486.

Price, J. L., & Morris, J. C. (1999). Tangles and plaques in nondemented aging and "preclinical" Alzheimer's disease. *Annals of Neurology, 45*(3), 358-368.

Qiu, C., Karp, A., von Strauss, E., Winblad, B., Fratiglioni, L., & Bellander, T. (2003). Lifetime principal occupation and risk of Alzheimer's disease in the Kungsholmen project. *American Journal of Industrial Medicine, 43*(2), 204-211.

Raber, J., Huang, Y., & Ashford, J. W. (2004). ApoE genotype accounts for the vast majority of AD risk and AD pathology. *Neurobiology of Aging, 25*(5), 641-650.

Rapoport, M., Dawson, H. N., Binder, L. I., Vitek, M. P., & Ferreira, A. (2002). Tau is essential to β-amyloid-induced neurotoxicity. *Proceedings of the National Academy of Sciences of the United States of America, 99*(9), 6364-6369.

Richards, M., & Dreary, I. J. (2005). A life course approach to cognitive reserve: A model for cognitive aging and development? *Annals of Neurology, 58*, 617-622.

Richards, M., & Sacker, A. (2003). Lifetime antecedents of cognitive reserve. *Journal of Clinical and Experimental Neuropsychology, 25*(5), 614-24.

Riley, K. P., Snowdon, D. A., Desrosiers, M. F., & Markesbery, W. R. (2005). Early life linguistic ability, late life function, and neuropathology: Findings from the Nun Study. *Neurobiology of Aging, 26*(3), 341-7.

Riley, K. P., Snowdon, D. A., & Markesbery, W. R. (2002). Alzheimer's neurofibrillary pathology and the spectrum of cognitive function: Findings from the Nun Study. *Annals of Neurology, 51*(5), 567-577.

Ritchie, K., Kildea, D., & Robine, J. (1992). The relationship between age and the prevalence of senile dementia: A meta-analysis of recent data. *International Journal of Epidemiology, 21*(4), 763-769.

Roberson, E. D., Scearce-Levie, K., Palop, J. J., Yan, F., Cheng, I. H., Wu, T., et al. (2007). Reducing endogenous tau ameliorates amyloid β-induced deficits in an Alzheimer's disease mouse model. *Science, 316*(5825), 750-754.

Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., et al. (2000). Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. *Applied Neuropsychology, 7*(1), 17-24.

Ruitenberg, A., van Swieten, J. C., Witteman, J. C., Mehta, K. M., van Duijn, C. M., Hofman, A., et al. (2002). Alcohol consumption and risk of dementia: The Rotterdam study. *Lancet, 359*(9303), 281-286.

Rylett, R. J., Ball, M. J., & Colhoun, E. H. (1983). Evidence for high affinity choline transport in synaptosomes prepared from hippocampus and neocortex of patients with Alzheimer's disease. *Brain Research, 289*, 169-175.

SAS Institute Inc. (2009). *SAS/STAT (R) 9.2 user's guide, 2nd ed.* Cary, NC: SAS Institute Inc.

Saunders, A., Hulette, C., Welsh-Bohmer, K., Schmechel, D., Crain, B., Burke, J., et al. (1996). Specificity, sensitivity, and predictive value of apolipoprotein-E genotyping for sporadic Alzheimer's disease. *The Lancet, 348*(9020), 90-93.

Scarmeas, N., & Stern, Y. (2003). Cognitive reserve and lifestyle. *Journal of Clinical and Experimental Neuropsychology, 25*(5), 625.

Scarmeas, N., & Stern, Y. (2004). Cognitive reserve: Implications for diagnosis and prevention of Alzheimer's disease. *Current Neurology and Neuroscience Reports, 4*, 374-380.

Schellenberg, G., & Turcotte, M. (2007). *A portrait of seniors in Canada.* Ottawa, ON: Statistics Canada. Retrieved from: http://www.statcan.gc.ca/pub/89-519-x/89-519-x2006001-eng.pdf

Scheuner, D., Eckman, C., Jensen, M., Song, X., Citron, M., Suzuki, M., et al. (1996). Secreted amyloid beta -protein similar to that in the senile plaques of Alzheimer's disease is increased in vivo by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease. *Nature Medicine, 2*(8), 864-870.

Schmand, B., Smit, J., Lindeboom, J., Smits, C., Hooijer, C., Jonker, C., et al. (1997). Low education is a genuine risk factor for accelerated memory decline and dementia. *Journal of Clinical Epidemiology, 50*(9), 1025-33.

School Sisters of Notre Dame. *Home page.* Retrieved 05/14, 2011, from www.ssnd.org

Segalowitz, N. S. (1981). Issues in the cross-cultural study of bilingual development. In H. C. Triandis, & A. Heron (Eds.), *Handbook of cross-cultural psychology: Developmental psychology*. Boston: Allyn & Bacon.

Segalowitz, N. S. (1990). Asian-American educational achievements: A phenomenon in search of an explanation. *American Psychologist, 45*, 913-910.

Sims, N. R., Smith, C. C. T., Davison, A. N., Bowen, D. M., Flack, R. H. A., Snowden, J. S., et al. (1980). Glucose metabolism and acetylcholine synthesis in relation to neuronal activity in Alzheimer's disease. *The Lancet, 315*(8164), 333-336.

Skoog, I., & Gustafson, G. (2006). Update on hypertension and Alzheimer's disease. *Neurological Research, 28*(6), 605-611.

Smetanin, P., Kobak, P., Briante, C., Stiff, D., Sherman, G., & Ahmad, S. (2009). *Rising tide: The impact of dementia in Canada 2008 to 2038.* Toronto, ON: Alzheimer Society of

Canada. Retrieved from http://www.alzheimer.ca/docs/RisingTide/Rising%20Tide_Full%20Report_Eng_FINAL_secured%20version.pdf

Snowdon, D. A. (1997). Aging and Alzheimer's disease: Lessons from the Nun Study. *Gerontologist, 37*, 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*, 34-8.

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), 528-32.

Solé-Padullés, C., Bartrés-Faz, D., Junqué, C., Vendrell, P., Rami, L., Clemente, I. C., et al. (2009). Brain structure and function related to cognitive reserve variables in normal aging, mild cognitive impairment and Alzheimer's disease. *Neurobiology of Aging, 30*, 1114-1124.

St George-Hyslop, P., Tanzi, R., Polinsky, R., Haines, J., Nee, L., Watkins, P., et al. (1987). The genetic defect causing familial Alzheimer's disease maps on chromosome 21. *Science, 235*(4791), 885-890.

Staff, R. T., Murray, A. D., Deary, I. J., & Whalley, L. J. (2004). What provides cerebral reserve? *Brain, 27*, 1191-1199.

Stelzmann, R. A., Schnitzlein, H. N., & Murtagh, F. R. (1995). An English translation of Alzheimer's 1907 paper "uber eine eigenartige erkankung der hirnrinde". *Clinical Anatomy, 8,* 429-431.

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society, 8*, 448-60.

Stern, Y. (2006). Cognitive reserve and Alzheimer disease. *Alzheimer Disease and Associated Disorders, 20*(2), 112-117.

Stern, Y. (2009). Cognitive reserve. *Neuropsychologia, 47*, 2015-28.

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), 1004-10.

Stern, Y., Tang, M. X., Denaro, J., & Mayeux, R. (1995). Increased risk of mortality in Alzheimer's disease patients with more advanced educational and occupational attainment. *Annals of Neurology, 37*, 590-5.

Sultzer, D., Levin, H., Mahler, M., High, W., & Cummings, J. (1993). A comparison of psychiatric symptoms in vascular dementia and Alzheimer's disease. *American Journal of Psychiatry, 150*(12), 1806-1812.

Swaab, D. F., Dubelaar, E. J. G., Hofman, M. A., Scherder, E. J. A., van Someren, E. J. W., & Verwer, R. W. H. (2002). Brain aging and Alzheimer's disease: Use it or lose it. In M.A.

Hofman, G.J. Boer, A.J.G.D. Holtmaat, E.J.W. van Someren, J. Verhaagen, D.F.Swaab (Eds.), *Progress in brain research* (pp. 343-373). Amsterdam, NL: Elsevier.

Swanwick, G., & Lawlor, B. A. (1999). Is female gender a risk factor for Alzheimer's disease? *International Psychogeriatrics,11*(3), 219-222.

Tedde, A., Nacmias, B., Ciantelli, M., Forleo, P., Cellini, E., Bagnoli, S., et al. (2003). Identification of new presenilin gene mutations in early-onset familial Alzheimer disease. *Archives of Neurology, 60*(11), 1541-1544.

Tyas, S. L. (2001). Alcohol use and the risk of developing Alzheimer's disease. *Alcohol Research and Health, 25*(4), 299-306.

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., Salazar, J. C., Snowdon, D. A., Desrosiers, M. F., Riley, K. P., Mendiondo, M. S., et al. (2007). Transitions to mild cognitive impairments, dementia, and death: Findings from the Nun Study. *American Journal of Epidemiology, 165*(11), 1231-1238.

Tyas, S.L., Snowdon, D.A., Desrosiers, M.F., Riley, K.P., & Markesbery, W.R. (2007). Healthy ageing in the Nun Study: Definition and neuropathologic correlates. *Age and Ageing, 36*(6), 650-655.

Tyas, S. L., Snowdon, D. A., Desrosiers, M. F., Riley, K. P., & Markesbery, W. R. (2009). Early-life linguistic ability, late-life pathology and asymptomatic Alzheimer's disease: findings from the Nun Study. *Alzheimer's and Dementia, 5*(4), S103.

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

Valenzuela, M. J., & Sachdev, P. (2005). Brain reserve and dementia: A systematic review. *Psychological Medicine, 36*, 441-454.

Valenzuela, M. J., & Sachdev, P. (2006). Brain reserve and cognitive decline: A non-parametric systematic review. *Psychological Medicine, 36*, 1065-1073.

Welsh, K. A., Butters, N., Hughes, J. P., Mohs, R. C., & Heyman, A. (1992). Detection and staging of dementia in Alzheimer's disease: Use of the neuropsychological measures developed for the Consortium to Establish a Registry for Alzheimer's disease. *Archives of Neurology, 49*(5), 448-452.

White, L., Katzman, R., Losonczy, K., Salive, M., Wallace, R., Berkman, L., et al. (1994). Association of education with incidence of cognitive impairment in three established populations for epidemiologic studies of the elderly. *Journal of Clinical Epidemiology, 47*(4), 363-374.

Whitehouse, P., Price, D., Struble, R., Clark, A., Coyle, J., & Delon, M. (1982). Alzheimer's disease and senile dementia: Loss of neurons in the basal forebrain. *Science, 215*(4537), 1237-1239.

Wilson, R. S., Barnes, L. L., & Bennett, D. A. (2003). Assessment of lifetime participation in cognitively stimulating activities. *Journal of Clinical and Experimental Neuropsychology, 5*(634), 642.

Wimo, A., & Prince, M. (2010). *World Alzheimer Report 2010: the global economic impact of dementia.* London, UK: Alzheimer's Disease International. Retreived from: http://www.alz.org/documents/national/World_Alzheimer_Report_2010.pdf

Wyss-Coray, T. (2006). Inflammation in Alzheimer disease: Driving force, bystander or beneficial response? *Nature Medicine, 12*, 1005-1015.

Appendices

Appendix A: Articles reviewed in literature search concerning multilingualism and cognition

Table 1. Summary of methods and samples used in studies of multilingualism and cognition.

Study	Study design; study question	Exposure	Outcome	Sample characteristics
Bialystok et al. 2008	- XS - what are the effects of bilingualism on working memory, executive control, and lexical fluency? - explored lifespan trajectories - hypothesized that bilinguals outperform monolinguals in executive control; bilinguals worse than monolinguals in lexical access	- bilingual vs. monolingual - self-report questionnaire was used to assess bilingualism - 45 claimed English as only language; 18 claimed Spanish as their only language; 19 considered themselves Spanish-English bilinguals - self-identified bilinguals were further assessed using a 20-item bilingual questionnaire asking the age at which the language was acquired and in what manner, how much contact with both languages was acquired, and where/when languages were used; a self-rated questionnaire about how proficient they were in each language; BNT in Spanish and English - variety of 2nd languages: most popular were French (n=7), Polish (n=7), Cantonese (n=6) and Spanish (n=4)	- Outcomes: Verbal fluency (VF), executive control (EC) and working memory (WM) - VF score assessed using: 1) phonetic VF: 3 x 1-min trials of words beginning with F, A, and S 2) semantic VF: 2 x 1-min trials of animals and fruits 3) Oral description of a picture 4) Sentence repetition (14 sentences, had to repeat as many as possible) - WM assessed using Corsi block test - EC measured using Simon Task	- n=96 (24 younger monolinguals (mean age = 20.7 yrs), 24 young bilinguals (mean age = 19.7), 24 older monolinguals (mean age = 67.2) and 24 older bilinguals (mean age = 68.3 yrs) - younger participants recruited from undergrad psych research pool (all attended university and had complete education in English, received compensation in the form of marks) - older participants were volunteers from a participant pool and received monetary compensation - young bilinguals: 14 were immigrants but had been in Canada before age 6 and all had formal education in English - older bilinguals: 20 were immigrants; all had arrived before age 20; (16 before age 12)

Table 1. Summary of methods and samples used in studies of multilingualism and cognition.

Study	Study design; study question	Exposure	Outcome	Sample characteristics
Fernandes et al. 2007	- XS - would semantic relatedness between memory and distracting task increase memory decrement associated with distracted attention (DA) in memory encoding and retrieval? - would younger adults have better recall under DA than older adults? - would bilinguals have better recall under DA than monolinguals?	- bilinguals vs. monolinguals - all language information ascertained using a self-report language questionnaire - variety of languages represented (Cantonese, Italian, Portuguese, Hindi, French, Spanish, Greek, Hebrew, Arabic) with no one language overrepresented - bilinguals used both languages on a daily basis - older bilinguals required to use two languages since age 12; younger bilinguals required to use two languages since age 6	- Outcome was recall, accuracy, and response times as mediated by EC - participants instructed to remember words from a 20-word list but not told words from a specific category - then prompted to repeat all words after 30 seconds - five conditions: full attention in encoding (memorization) and retrieval (recall); DA from related words in encoding; DA from unrelated words in encoding; DA from related words in retrieval; DA from unrelated words in retrieval - accuracy of recall measured, number of words recalled correctly measured, response time measured, and % of recall from full attention measured	- n=104; 52 young undergraduate students (mean age = 20.5 years; 36 females and 16 males) recruited through class - 52 older adults (mean age = 70.1 years; 36 female and 16 male) were recruited via posters in the community - in each group, half were monolinguals and half were bilinguals

Table 1. Summary of methods and samples used in studies of multilingualism and cognition.

Study	Study design; study question	Exposure	Outcome	Sample characteristics
Craik et al. 2006	- What is the effect of age on cognitive planning? - What are the effects of bilingualism on planning behaviour? -prospective memory and WM had been known to suffer age-related losses	- bilingual vs. monolingual - language background questionnaire was used to assess exposure status - contained questions about # of languages spoken, where they were learned, the means by which they were learned, how often they are used in daily life, and how often they are used at present day - included self-rating scale to determine proficiency in understanding, reading, writing and speaking in each language - bilinguals were those who had spoken 2 languages every day from an early age (younger: <6 yrs old; older: <10 yrs old) and continued to use them up to the present with excellent proficiency - bilingual participants had a variety of second languages	- outcome was performance on task of "cooking" five foods by clicking on relevant computer icons, monitor the progress and stop foods once they were done - participants also had to "set the table" as a distracter task - task measured perseveration (suspend table setting in order to check/cook food) as well as prospective memory (to remember to start and stop foods accordingly) and WM (hold progress of foods and general plan in mind) - task was presented in 3 levels of ascending difficulty - measured discrepancy between desired end time and actual time = measure of prospective memory - looked at range of stop times; this reflected planning ability + WM - looked at average deviation of start times to measure WM	- 60 participants were tested; half between ages 18 and 30 (mean=20.2) and half were older between ages 60 and 80 (mean=69.6) - In each age group, half were monolingual and half were bilingual - 4 groups (young bilingual, young monolingual, older bilingual, older monolingual) were matched based on years of education - general background questionnaire was used to establish the age, education level, and health status of each participant (included medications) - no information regarding how participants were selected from the general population

Table 1. Summary of methods and samples used in studies of multilingualism and cognition.

Study	Study design; study question	Exposure	Outcome	Sample characteristics
Bialystok et al. 2006	- XS - Do bilinguals have advantages over monolinguals in EC tasks that do not involve language? - used 2 experiments: 1) using anti-saccade task 2) using key-press instead of anti-saccades - hypothesized bilinguals would have advantages for inhibitory control - hypothesized older bilinguals would have advantages over the older monolinguals	- bilingual vs. monolingual - used a language history questionnaire similar to those used in prior studies (i.e. Craik et al. 2006) to assess bilingualism - All bilinguals reported using both English and the other language daily; rated proficiency in both English and the non-English language on a 5-point scale marked from poor (0) to excellent (4) - mean rating for speaking ability by the young bilinguals was 3.83 (SD 0.39) for English and 3.15 (SD 0.90) for the other language - corresponding results for the older bilinguals were 3.79 (SD 0.41) for English and 3.65 (SD 0.57) for the other language	- outcomes were WM, lexical fluency and EC - WM: assessed using forward and backward Corsi block span and self-ordered pointing task - lexical fluency: measured using Peabody picture vocabulary test, BNT, letter VF and category VF - EC: measured with Simon task, Stroop task, and the sustained attention to response task	**Experiment 1:** n = 24 in each category of young monolingual, young bilingual, older monolingual and older bilingual - young participants were volunteers from undergraduate psychology classes and older participants were from a research pool - young monolinguals had a mean age = 20.7 yrs w/14.4 yrs education; young bilinguals average age = 20.8 yrs w/14.6 yrs education - old monolinguals average age = 70.4 yrs w/ 15.5 years education; old bilinguals average age =71.3 yrs w/ 16.6 yrs education **Experiment 2:** a "new" group of 96; 24 in each group, were recruited from similar populations as study 1 - used same methods of exposure assessment; young monolinguals mean age = 25.6 years, 16.8 yrs education, bilinguals mean age=23.9 yrs w/ 16.5 yrs education - older monolinguals mean age =66.9 yrs w/15.3 education, older bilinguals mean age =64.5 yrs with 14.5 yrs education

Table 1. Summary of methods and samples used in studies of multilingualism and cognition.

Study	Study design; study question	Exposure	Outcome	Sample characteristics
Bialystok et al. 2004	- XS - Are executive processing advantages in bilingual children also found in bilingual adults? - Are advantages maintained in older age, perhaps protecting adult bilinguals from the normal decline that occurs with age?	- bilingual vs. monolingual - exposure ascertained via self-report questionnaire (similar to that used by Bialystok et al. in previous studies) - bilinguals reported using both English and the other non-English language daily and could understand, speak, write, read in other language with excellent proficiency - younger bilinguals had to be using both languages since age 6; older since age 12	- outcome was EC and recall - EC measured using two relevant studies of the Simon task - **Study 1**: Simon task between older and younger adults, bilinguals and monolinguals **- Study 2**: more trials than first, two more colours added to interference - recall measured by # of words recalled in the presence of a distracter - participants asked to recall them orally in any order; done either with interference from another list of words (presented visually) or without - 5 conditions: no distracters; distracters (related to test words) during memorization/"encoding" process; distracters (unrelated) during memorization; distracters (related) during recall; distracters (unrelated) during recall	**Study 1:** 40 participants composed two language groups and two age groups: 20 were 30-54 years old and 20 were older (60-88 yrs) - in each age group, half were English-speaking monolinguals living in Canada; rest were Tamil-English speaking bilinguals living in India; age matched - equal numbers of genders - same experimenter, procedures - younger participants recruited via email; older participants via flyers **Study 2:** 94 participants; younger adults had 64 participants, ranging from 30-58 years (mean=42.6) divided evenly into Canadian English monolinguals and Eng-Tamil bilinguals from India (n=20) or Cantonese-English participants in Hong Kong (n=12); each bilingual age matched to a monolingual - equal genders in each group - older group: n=30 participants 60-80 years old divided between English monolinguals and English-Tamil (n=9) or English-French living in Canada (n=6) - same recruitment methods

Table 1. Summary of methods and samples used in studies of multilingualism and cognition.

Study	Study design; study question	Exposure	Outcome	Sample characteristics
Rosselli et al. 2000	- XS - What is the impact of bilingualism on VF in older Hispanic persons?	- bilingualism vs. monolingualism - exposure ascertained by language questionnaire similar to that used by the Bialystok papers - used both languages on a daily basis, can read, write, understand, and speak with excellent proficiency; learned both languages before age 6 for younger participants and 12 for older participants	- 4 outcomes were used: 1. phonemic VF (using letters F, A, and S) 2. semantic VF (fruits and animals) 3. oral description of the picture *The Cookie Theft* (descriptions were recorded using a tape recorder) → score was the total # of words 4. Sentence repetition from both the Spanish and English versions of the Multilingual Aphasia Examination (MAE) → recorded the total # of correctly repeated sentences	- n=82; 28 men, 54 women; 19 bilinguals and 63 monolinguals - all bilinguals claimed Spanish as their 1st language (L1) - all were residents from South Florida who volunteered to be in the study, claimed to be Spanish or English speaking monolinguals or Span-English bilinguals - Were screened for any psychiatric or neurological problems before entry into study: all lived independently and could complete ADLs; MMSE, Beck depression inventory were used to determine cognitive health; all were non-depressed (<5 on depression inventory) and scored 27 < on the MMSE; BNT was used to test naming proficiency and participants yielding abnormal results were excluded - Spanish-only monolinguals were tested with Spanish versions of the previous tests

Abbreviations: BNT = Boston Naming Test; EC = executive control; MMSE = Mini-mental state examination; PC = prospective cohort; RT = reaction time; VF = verbal fluency; WM = working memory; XS = cross-sectional; DA = divided attention

Table 2. Summary of findings from studies of multilingualism and cognition.

Study	Covariates and descriptive stats	Statistical tests used/confounders addressed	Principal findings
Bialystok et al. 2008	- covariates include education, age - no difference in education between the young bilinguals and monolinguals (12.36 yrs (SD 0.95) and 12.83 yrs (SD 1.30), respectively) - no difference in educational years between older bilinguals and monolinguals (14.25 yrs (SD 2.45) and 14.43 yrs (SD 1.43), respectively) - older participants had significantly more education than the younger participants ($p<0.0001$) - this was not surprising since the younger adults weren't old enough to attain as much education as the older participants	- WM**:** 2-way ANOVAs analysed effects of exposure and age on the Corsi test - VF: 2-way ANOVAs analysed effects of exposure and age group on the PPVT and BNT - EC: Simon error rate was negligible - 2-way ANOVA for age group and language group was done for RTs only - 3-way ANOVA for Simon outcome RTs stratified by age, exposure group and conditions - 3-way ANOVA for Simon RTs stratified by condition, age and exposure groups - 2-way ANOVA for % increase of Simon Task RTs depending on condition, stratifying by age, and exposure group - ANOVAs performed for Stroop RTs by age, exposure group, and condition (colour or word naming)	- **WM:** interaction between age and language group: the younger bilinguals recalled more blocks than the younger monolinguals - older bilinguals' performances didn't differ significantly from the older monolinguals - the backwards (harder) block test took longer for the older vs. younger participants but language groups did not significantly differ - **VF:** monolinguals outperformed bilinguals in both tasks ($p<.0001$ for each) - age had no influence on performance; no significant interactions present - **EC: -** no effect of age or languages spoken - after stratifying by stimulus direction, only age was significant (younger outperformed older) - interaction between congruence, age and exposure group; the differences in RTs between the congruent and incongruent trials were not significant (they were more similar) in the older bilinguals, while it was significantly different for all other groups - the % increase in RT was significantly larger for monolinguals in the older category, while it wasn't significantly larger for any other group (young monolinguals or bilinguals, older bilinguals) - larger Stroop effect (difference in RTs between congruent and incongruent trials) for monolinguals compared to bilinguals in both age groups

Table 2. Summary of findings from studies of multilingualism and cognition.

Study	Covariates and descriptive stats	Statistical tests used/confounders addressed	Principal findings
Fernandes et al. 2007	- covariates were education, age (older or younger), divided attention condition, and category of word relatedness - no significant differences in education between monolinguals and bilinguals - older participants had significantly more years of education than the younger participants	- 4-way ANOVAs were conducted for outcomes from distracted attention tasks; stratified by age, exposure group, phase of attention measured, and relatedness of categories - hierarchical linear regression evaluated changes in R^2 for each predictor; these used age, PPVT and Cattell raw score as covariates in model - ANOVA stratified by age and exposure groups to look at influence on total word recall - used 4-way ANOVA, stratified by age, and phase of attention in order to assess outcome of % recall decline between exposure groups	- in distracted attention tasks, younger participants significantly outperformed older participants (expected result) but monolinguals significantly outperformed bilinguals in each age group (unexpected result) - older monolinguals recalled more words than older bilinguals (significantly more in only the full attention task) - older monolinguals experienced a less drop in words recalled when looking at % of words recalled from full attention conditions (unexpected result); no effect of language was significant - older monolinguals had higher accuracies and faster response times than older bilinguals (unexpected result) under full attention; no differences when distracters were introduced - no interactions were found between age group and bilingualism - it is possible that bilinguals have smaller vocabularies since they use two languages as often as monolinguals use one; this may have influenced the findings

Table 2. Summary of findings from studies of multilingualism and cognition.

Study	Covariates and descriptive stats	Statistical tests used/confounders addressed	Principal findings
Craik et al. 2006	- covariates were age category and the difficulty level of the Cooking Breakfast task	- 3-way ANOVA tests stratifying by age group, exposure group and task difficultly level were used to evaluate influence on outcome	Effects of age differences: - substantial age effects (older participants had poorer performances) as the situations became harder - younger adults set more places - older adults spent more time setting inappropriately, had significantly higher discrepancy times and significantly larger ranges of stop times - discrepancies between desired and actual end time were lower for the younger participants ($p<.0001$) with no differences between exposure groups and no interactions Effects of bilingualism: - smaller/more subtle than effects of aging - in both age groups, the monolinguals and bilinguals performed similarly on all the tasks reflecting prospective memory and WM - bilingual advantage on the place setting task = bilinguals spent less time place setting, less inappropriate time spent setting when should have been checking/cooking food compared to monolinguals - results suggest bilinguals are more effective in switching tasks or ignoring distracting stimuli

Table 2. Summary of findings from studies of multilingualism and cognition.

Study	Covariates and descriptive stats	Statistical tests used/confounders addressed	Principal findings
Bialystok et al. 2006	- covariates were age category, and task presentation (mixed or blocks of saccades)	**Study 1:** not specified which analyses were used; assumed 3-way ANOVA since interactions between language group, age group and condition of task were reported - 4-way ANOVA for RTs stratified by age, exposure, task, and presentation **Study 2:** 4-way ANOVA for age, exposure, task, and presentation - 2-way ANOVA looking at processing costs of EC, stratifying by age and exposure	**Study 1:** - no significant effects of bilingualism or interactions for RTs - no effects of bilingualism or interactions with language were found for any analyses **Study 2:** bilinguals had smaller RTs than monolinguals ($p<.0009$) with an interaction between age and exposure, meaning younger bilinguals didn't perform any different while the older bilinguals performed significantly faster than the older monolinguals - incongruent trials (harder tasks) were more costly (RTs larger) for monolinguals than for bilinguals ($p<.0001$) - significant effect of bilingualism on processing costs; ($p<.0005$) and interaction between age and exposure was significant ($p<.01$) for both processing costs and task switching, meaning bilingual advantage was found in only older adults

Table 2. Summary of findings from studies of multilingualism and cognition.

Study	Covariates and descriptive stats	Statistical tests used/confounders addressed	Principal findings
Bialystok et al. 2004	- covariates were age category, education, SES - all participants had bachelor's degrees and similar middle class SES backgrounds, although it was not specified how this information was ascertained **Study 1:** baseline scores were no different between the age and language groups **Study 2:** same covariates as 1st study - groups did not differ across baseline measures	**Study 1:** compared mean Simon task accuracy scores using a 3-way ANOVA, stratifying by age group (younger/older), exposure group and trial congruency - RTs from Simon task were compared using a 3-way ANOVA for same variables **Study 2:** used same 3-way ANOVA for testing the errors in Simon task - 4-way ANOVA examined the RT from Simon task by stratifying by age group, exposure group, color of stimulus, and congruency	**Study 1:** - interactions were present between exposure group and congruency of trial ($p<.01$) and between exposure, age, and congruency ($p<.01$), meaning older monolinguals in the incongruent trial had more errors - bilinguals were significantly faster than monolinguals - older adults and monolingual adults had more difficulties on the incongruent trials but there was no interaction between age, exposure group and congruency (i.e. the increase in reaction time for incongruent in older adults was the same regardless of language group) **Study 2:** younger adults had more errors than older participants ($p<.01$) - no difference in errors between exposure groups - significant interaction between exposure group and age group, therefore the higher accuracy rate for the older participants was stronger in the bilinguals - with age, the monolinguals had larger increases in RTs than the bilinguals, therefore larger WM costs in older monolinguals than bilinguals

Table 2. Summary of findings from studies of multilingualism and cognition.

Study	Covariates and descriptive stats	Statistical tests used/confounders addressed	Principal findings
Rosselli et al. 2000	- age of acquisition for the L2 was the only covariate analysed - mean age of sample = 61.76 (SD 9.30) - mean education = 14.8 years (SD 3.6) - 47% (n=9) participants had contact with English before age 12; 53% (n=10) after age 12 - mean age of exposure to L2 was 18.85 yrs (SD 14.24) - mean number of years exposed to English was 35.95 (SD 13.37) yrs - All but one bilingual participant were immigrants to the US - 63% of bilinguals spoke mostly Spanish at home - no significant differences between exposure groups in age, level of education, and MMSE	- ANOVA tests were used to determine significant differences in outcomes between exposure groups - paired t-tests were used to determine any differences between languages used by the bilinguals - 2-way ANOVA was used to measure the effect of age of L2 acquisition on performance	- bilinguals took a mean time of 75 mins. to complete the VF tests; monolinguals took mean time of 45 mins. - bilinguals produced significantly less words in both English (fruit and animal) and Spanish (fruit) than the monolinguals - no significant differences in amounts of words produced in the phonemic (letter) categories between bilinguals and monolinguals - bilinguals generated more words from the picture description task in English rather than in Spanish even though more claimed they had Spanish as their L1 - age of acquisition was a effect modifier; there were significantly more words generated in the mother tongue if the L2 was acquired after the age of 12

* ANCOVA = analysis of covariance; ANOVA = analysis of variance; Cattell = Cattell Culture Intelligence Test; EC = executive control; L1 =first language acquired; L2 = second language acquired; LSD = least significant difference; PPVT = Peabody Picture Vocabulary Test; RT = reaction time; VF = verbal fluency WM = working memory

Table 3. Summary of methods used in studies examining multilingualism and cognitive decline.

Study	Study design; study question	Exposure	Outcome	Sample characteristics
Craik et al. 2010	- XS - Follow-up to Bialystok 2007 - Does bilingualism confer additional CR, with regard to the delay of AD onset?	- bilingualism, as defined as regularly using at least 2 languages for the majority of life - classified 102 patients as bilingual, (60 female) and 109 monolingual (60 female) - bilinguals included speakers of 21 first languages, of which the most common were Yiddish (n=24), Polish (n=12), Italian (n=11), Hungarian (n=9), and French (n=7)	- 2 outcomes: age of AD symptom onset and age of 1st clinic appointment - At 1st clinic visit, age at onset of cognitive impairment was recorded from patients or caregivers	- n=211 probable AD patients from Baycrest memory clinic in Toronto; diagnosed using NINCDS-ADRDA criterion between January 2007 and December 2009
Chertkow et al. 2010	- XS - Do multilinguals have prolonged periods of time before AD onset?	- multilingualism; classified as monolingual, bilingual, trilingual, or multilingual (4+ languages) - divided sample by those who said that they were most fluent in their mother tongue (n=366; 45%) or those who were most fluent in second/third/fourth language (n=448; 55%)	- Outcomes were scores on 2 cognitive screening tests: - MMSE and the cognitive screening test developed by Katzman et al. (1983) which measures time orientation, memory and concentration	- Based on information from the memory clinic database at the Jewish General hospital in downtown Montreal - 1842 participants were referred between 1997 and 2006; restricted to only participants subsequently diagnosed with AD (n=632)

Table 3. Summary of methods used in studies examining multilingualism and cognitive decline.

Study	Study design; study question	Exposure	Outcome	Sample characteristics
Kavé et al. 2008	- XS - Does the number of languages spoken predict performance on cognitive screening tests?	- multilingualism; classified as monolingual, bilingual and multilingual (3+ languages) - n= 253 multilinguals (53% immigrants) who were further divided into: 168 bilinguals, 67 trilinguals and 18 who spoke 4+ languages - 379 monolinguals (77% only English and 23% only French) - Language history was obtained from patient and caregiver interviews	- 2 outcomes: age of AD diagnosis and age of AD symptom onset (for a subset of participants) - diagnosis information ascertained by using medical records - AD diagnoses made using NINCDS-ADRDA criterion - 143 participants had further info about age of symptom onset from family interviews ("can you give the month and year when you first noticed memory problems in the patient")	- Random sample of an older Jewish Israeli population, using the National Population Registry on Jan. 1, 1989 - initial sample was 2400; for this study 15.7% of original sample had died or were not located before the sampling day and 8.5% refused to be interviewed. - 1820 (75.8%) interviews were conducted in wave 1; of this 75.2% were in-person interviews - all interviews were conducted in home by multilingual interviewers - n = 814 were in wave 1 (1989), n = 457 in wave 2 (on average, 3.5 years later, approx. 1992) and n = 115 in wave 3 (8.2 years after 2nd wave)

Table 3. Summary of methods used in studies examining multilingualism and cognitive decline.

Study	Study design; study question	Exposure	Outcome	Sample characteristics
Bialystok et al. 2007	- XS - Does bilingualism contribute to CR and protect older adults from cognitive decline?	- bilingualism vs. monolingualism - criterion for bilingualism: patients had to have spent most of their lives speaking both languages - inter-rater reliability was .95 for monolingual assignment and .81 for bilingual - only patients schooled 5+ years in English and used both languages at work for 10+ years were selected	- AD as diagnosed by the NINCDS-ADRDA criterion; other forms of dementia or CVD - outcomes were ascertained through medical records - proxies were asked when exactly the first AD symptoms were evident in the cases	- examined records of 228 patients at Baycrest Memory clinic (downtown Toronto) between 2002-2005 with memory complaints - records included medical history, physical exam, CT, SPECT, blood test - 44 patients who received a psychiatric diagnosis other than dementia or couldn't be classified w/exposure were excluded - final sample had 184 patients; 91 were monolingual and 93 bilingual; 66 in each exposure group were diagnosed with AD - 52 were diagnosed with other kinds of dementia or CVD

* AD = Alzheimer disease; CT and SPECT = types of medical imaging; CVD = cerebrovascular disease; CR = cognitive reserve; MMSE = Mini-mental state examination; NINCDS-ADRDA = National Institute of Neurological and Communicative Disorders and Stroke – Alzheimer's Disease and Related Disorders Association; XS = cross-sectional

Table 4. Summary of analyses and findings from studies of multilingualism and cognitive decline.

Study	Covariates and descriptive stats	Statistical tests used	Principal findings
Craik et al. 2010	- covariates included gender, occupational history, education, MMSE scores, and date of immigration to Canada - occupation data was based on 4-point scale developed by Human Resources and Skills Development, Canada, in which higher numbers signify higher status - monolingual participants reported more formal schooling than bilingual participants (p <0.003)	- 2-way ANOVA for lingual status and gender, comparing ages of onset - effect of immigration was compared through 2-way ANOVAs (language group x immigration status) on the 2 age variables	- monolinguals had a 5.1-year younger age of onset compared to bilinguals ($p < 0.0001$) - monolinguals presented at clinic 4.3-years earlier than bilinguals ($p< 0.0006$) - bilingual group contained more immigrants than the monolingual group - duration between symptom onset and clinic visit was longer for monolingual participants (3.8 years) than bilingual participants (3.1 years) ($p< 0.05$) - monolingual group included 35 immigrants and 74 non-immigrants; bilingual group contained 81 immigrants and 21 non-immigrants - For age of symptom onset, ($p<0.0001$) and clinic presentation age ($p<0.0007$) there were effects of language group but no effect of immigrant status - no differences between language groups for MMSE scores or gender

Table 4. Summary of analyses and findings from studies of multilingualism and cognitive decline.

Study	Covariates and descriptive stats	Statistical tests used	Principal findings
Chertkow et al. 2010	- covariates included sex, education, and immigrant status; SES was mentioned yet not reported in results - immigrants had significantly less education than natives (9.7 vs 11.2 yrs, p<0.001)	- used cross-tabulations and Pearson chi-squared test to determine outcome differences between exposure groups - linear regression model to examine correlation coefficients - used post hoc LSD analysis to determine which groups were significantly different from one another (via pair-wise comparisons of variable means)	- No significant correlation was found between education and age at diagnosis - Multivariate regression analysis showed that the number of languages spoken had a small yet significant positive association with the age at diagnosis (Spearman coefficient = 0.14; p=0.026) - Those who spoke 4+ languages were diagnosed significantly later than those who spoke only 2 (mean difference = 4.19 years; p=0.02) - Those speaking 3+ languages had a later onset of symptoms than those who spoke only 1 (mean difference = 4.84 years; p=0.026) or 2 (mean difference = 5.47 years; p=0.0222) - Education and sex did not contribute significantly to the model - No significant differences between those who spoke one or 2 languages; the only benefit appeared to be when 3+ languages were spoken -Within the immigrant population, monolinguals were diagnosed with AD significantly earlier than bilinguals (5 years; p=.006), 6.4 years earlier than trilinguals (p=0.002) and 9.5 years earlier than those speaking 4+ languages - Bilingual immigrants were diagnosed 4.5 years earlier than the 4+ speakers (p=0.038) - Rate of cognitive decline did not alter according to number of languages spoken, education, gender or immigrant status

Table 4. Summary of analyses and findings from studies of multilingualism and cognitive decline.

Study	Covariates and descriptive stats	Statistical tests used	Principal findings
Kavé et al. 2008	- covariates: # years of education, age, place of birth, age at immigration - participants speaking more languages had higher average levels of education - 52.9% obtained <12 years of education and 36.2% had more than 12 years of education - Of Wave 1 Hebrew speakers, 86 (10.9%) individuals had no formal education (mean age 83.7 years, SD 5.1); women comprised 67% of the group	- ANCOVAs were performed for each outcome (all scores from Katzman test; only Wave 2 MMSE) - hierarchal regression was used to analyze Katzman test score prediction (using info from all waves) and MMSE (only wave 2) - looked at variable significance through values of R and β coefficients	- # of languages spoken had a significant effect on the cognitive scores across all 3 waves - β coefficients for age, gender and education level were significant in wave 1 and 2 for Katzman score prediction and MMSE scores - location of birth and age of immigration were not significant - "# of languages spoken" variable accounted for the most variance - 75-95 year olds speaking 4+ languages had best cognitive states, even up after age 90 and 12 years of follow up - differences between exposure groups remained significant across all waves in both tests, even when controlled for education - # of languages spoken also predicted higher cognitive function for those who had less than 12 years of education - those whose best language was their mother tongue scored worse on the Katzman test when compared to those whose best language was their $2^{nd}/3^{rd}/4^{th}$ language

Table 4. Summary of analyses and findings from studies of multilingualism and cognitive decline.

Study	Covariates and descriptive stats	Statistical tests used	Principal findings
Bialystok et al. 2007	- covariates were gender, years of education, age of symptom onset, immigrant status - occupational status was measured on a 5-point scale for a subset of participants (n=147)	- ANCOVAs were performed to evaluate outcomes by exposure group, while adjusting covariates - 2-way ANOVA was used to evaluate outcome, stratifying by immigrant status and language group - ANCOVA adjusting for occupational status was used to compare age of onset between exposure groups - linear regression was used to examine the decline in MMSE scores between exposure groups	- Monolinguals exhibited earlier onset of AD; $p<0.003$ (Mean age 71.4 (SD 9.6) vs. mean age 75.5 (SD 8.5)) - bilinguals were 3.2 years older than monolinguals at time of initial appointment ($p<0.02$) - men were more likely to wait to go to the memory clinic ($p<0.02$) - Bilinguals' delay of AD symptom onset was significantly greater compared to monolinguals; this was true both in the subsample of 132 patients with probable AD ($p<0.009$) with a delay of 4.3 years and for the other dementias ($p<0.04$) with a delay of 3.5 years - based on MMSE scores, the different language groups exhibited similar rates of cognitive decline over the 4 years → interesting since bilinguals had significantly less education than monolinguals (12.4 yrs of education (6.4) vs. 10.8 (4.2); $p<0.009$) - this educational difference might reflect differences in opportunity since most bilinguals were immigrants during WWII era - then looked at only monolinguals and bilinguals who were immigrants; the age of onset was significantly higher in bilinguals: 75.3 vs. 63.8 years; ($p<0.0001$)

* AD = Alzheimer disease; ANCOVA = analysis of covariance; ANOVA = analysis of variance; LSD = least significant difference; MMSE = mini-mental state examination

Appendix B: Descriptions of cognitive tests used in the studies of multilingualism and cognition

Test	Area of cognition	Test description
Corsi block test	Working memory	- Blocks are tapped or highlighted by an examiner or computer program in randomized sequences of increasing length - Immediately after each tapped sequence, the participant attempts to reproduce the sequence, progressing until no longer accurate - more or longer sequences = better scores
Peabody Picture Vocabulary Test	Verbal fluency	- involves a series of pictures being presented to the participant; each page has 4 numbered pictures. - the examiner states a word describing one of the pictures and asks the individual to point to or say the number of the picture that the word describes - higher accuracy = better scores
Boston Naming Test	Verbal fluency	- involves object naming from line drawings - test contains 60 items; they are rank ordered in terms of their ability to be named, which is thought to be correlated with their frequency in daily life - naming more objects = higher scores
Semantic verbal fluency test	Verbal fluency	- involves listing the most words relating to a certain category (e.g., 4-legged animals) as possible in a given time period (usually one minute) - more words = better scores - repeated words and words not relating to the category are excluded
Phonemic verbal fluency test	Verbal fluency	- involves listing the most words beginning with a certain letter (i.e. the letter F) as possible in a given time period (usually one minute) - more words = better scores - repeated words and words not relating to the category are excluded
Stroop task	Executive Control	- involves reciting the colour of words on a page, rather than reading the word itself (i.e. the word is "red" but it is in blue-coloured ink; in order to be correct one would say "blue")
Anti-saccade task	Executive control	- involves the tracking of participant eye movements in response to stimuli on a screen - pro-saccades are when participants look in the direction of the stimulus - anti-saccades are when participants look in the opposite direction of the stimulus - better performances = less errors, smaller reaction times

Test	Area of cognition	Test description
Simon Task	Executive control	- involves coloured stimuli on either sides of a screen, paired with response keys on a keyboard on the different sides - keys match to the stimuli - on congruent trials, the correct response for that colour is on the same side as the stimulus - on incongruent trials the correct response key is on the opposing side - increased time, going from congruent to incongruent trials to make the correct response is known as the "Simon effect" - usually observe larger Simon effect with older age

Appendix C. Tables of non-response comparisons

Table 1. Comparison of analytic sample to excluded participants and deceased participants (Research Question 1).

Variable	Total Nun Study population (n=678)	Analytic sample (n=157)	Excluded participants (n=521)[1]	Excluded participants that were dead (n=449)[2]
# of languages spoken[3] (%)				
1	29.6	28.7	30.0	29.0
2	50.5	50.3	50.6	51.7
3	15.6	17.2	14.9	15.9
4	2.4	1.27	2.3	2.1
5	2.0	2.55	1.7	1.4
Educational category (%)				
Grade school	10.0	7.0	10.9	11.6
High school	5.5	3.2	6.1	6.9
Bachelor's degree	39.8	40.8	39.6	40.1
Master's degree+	44.7	49.0	43.4	41.4

*p <0 .05; **p<0.01; ***p<0.001

[1] Tests of significance with this group were conducted in comparison to the analytic sample. Participants in this group include those missing information on key variables in addition to those excluded for not meeting the outcome criteria.

[2] Tests of significance with this group were conducted in comparison to the analytic sample. Participants in this group were any that were dead. This comparison was done since some of the participants originally excluded were those who were still alive (and would not have neuropathology information).

[3] In the total Nun Study population, n=507 had complete data concerning the # of languages spoken.

[4] In the total Nun Study population, n=619 had complete data concerning ApoE-E4 status.

[5] In the total Nun Study population, n=676 had complete data concerning occupation.

[6] With respect to participants in the total Nun Study population who were teachers (n=605).

[7] With respect to participants in the total Nun Study population who had handwritten autobiographies analysed for linguistic ability (n=180).

Table 1 (continued). Comparison of analytic sample to excluded participants and deceased participants (Research Question 1).

Variable	Total Nun Study population (n=678)	Analytic sample (n=157)	Excluded participants (n=521)[1]	Excluded participants that were dead (n=449)[2]
Immigrant to US (%)	6.1	4.5	6.5	6.7
% with an ApoE-E4 allele[4]	22.8	26.1	21.7	23.4
Age at death (mean)	90.4	90.9	90.2	90.2
Age at last cognitive assessment (mean)	89.5	90.2	89.3*	88.9**
Occupation[5]				
Teacher	89.5	93.0	88.4	87.7
House Sister	8.0	6.4	8.5	9.2
Nurse's Aide or Other	2.5	0.6	3.1	3.1
# of years teaching (mean)[6]	37.6	39.8	36.9*	36.7*

*p <0 .05; **p<0.01; ***p<0.001

[1] Tests of significance with this group were conducted in comparison to the analytic sample. Participants in this group include those missing information on key variables in addition to those excluded for not meeting the outcome criteria.

[2] Tests of significance with this group were conducted in comparison to the analytic sample. Participants in this group were any that were dead. This comparison was done since some of the participants originally excluded were those who were still alive (and would not have neuropathology information).

[3] In the total Nun Study population, n=507 had complete data concerning the # of languages spoken.

[4] In the total Nun Study population, n=619 had complete data concerning ApoE-E4 status.

[5] In the total Nun Study population, n=676 had complete data concerning occupation.

[6] With respect to participants in the total Nun Study population who were teachers (n=605).

[7] With respect to participants in the total Nun Study population who had handwritten autobiographies analysed for linguistic ability (n=180).

Table 1 (continued). Comparison of analytic sample to all excluded participants and deceased participants (Research Question 1).

Variable	Total Nun Study population (n=678)	Analytic sample (n=157)	Excluded participants (n=521)[1]	Excluded participants that were dead (n=449)[2]
Linguistic ability[7]				
Quartile of grammatical complexity ability (%)				
1 (low)	24.4	27.1	23.5	29.1
2	25.0	31.2	22.7	21.4
3	25.6	18.8	28.0	24.3
4 (high)	25.0	22.9	25.8	25.2
Quartile of idea density ability (%)				
1 (low)	25.0	25.0	25.0	30.1
2	25.6	20.8	27.3	25.2
3	23.9	27.1	22.7	19.4
4 (high)	25.6	27.1	25.0	25.2

*p <0 .05; **p<0.01; ***p<0.001

[1] Tests of significance with this group were conducted in comparison to the analytic sample. Participants in this group include those missing information on key variables in addition to those excluded for not meeting the outcome criteria.

[2] Tests of significance with this group were conducted in comparison to the analytic sample. Participants in this group were any that were dead (either excluded or in analytic sample)

[3] In the total Nun Study population, n=507 had complete data concerning the # of languages spoken.

[4] In the total Nun Study population, n=619 had complete data concerning ApoE-E4 status.

[5] In the total Nun Study population, n=676 had complete data concerning occupation.

[6] With respect to participants in the total Nun Study population who were teachers (n=605).

[7] With respect to participants in the total Nun Study population who had handwritten autobiographies analysed for linguistic ability (n=180).

Table 2. Comparison of analytic sample to excluded participants, separated by how they were excluded from the analytic sample (Research Question 1).

Variable	Total Nun Study population (n=678)	Analytic sample (n=157)	Excluded participants for missing data; (n=400)[1]	Excluded participants who did not meet criteria for outcomes; (n=121)[2]
# of languages spoken[3] (%)				
1	29.6	28.7	31.0	28.1
2	50.5	50.3	47.6	56.2
3	15.6	17.2	17.5	9.9
4	2.4	1.27	2.6	3.3
5	2.0	2.55	1.3	2.5
Educational category (%)				
Grade school	10.0	7.0	12.5	5.8
High school	5.5	3.2	5.8	7.4
Bachelor's degree	39.8	40.8	38.3	43.8
Master's degree+	44.7	49.0	43.5	43.0
Immigrant to US (%)	6.1	4.5	6.8	5.8
% with an ApoE-E4 allele[4]	22.8	26.1	21.7	21.5
Age at death (mean)	90.4	90.9	89.6*	91.9
Age at last cognitive assessment (mean)	89.5	90.2	88.7**	91.2

*p <0 .05; **p<0.01; ***p<0.001

[1] Tests of significance with this group were conducted in comparison to the analytic sample. Participants in this group include those missing information on key variables in addition to those excluded for not meeting the outcome criteria.

[2] Tests of significance with this group were conducted in comparison to the analytic sample. Participants in this group were any that had complete data but did not meet criteria for the outcomes of this investigation.

[3] In the total Nun Study population, n=507 had complete data concerning the # of languages spoken.

[4] In the total Nun Study population, n=619 had complete data concerning ApoE-E4 status.

[5] In the total Nun Study population, n=676 had complete data concerning occupation.

[6] With respect to participants in the total Nun Study population who were teachers (n=605).

[7] With respect to participants in the total Nun Study population who had handwritten autobiographies analysed for linguistic ability (n=180).

Table 2 (continued). Comparison of analytic sample to excluded participants, separated by how they were excluded from the analytic sample (Research Question1).

Variable	Total Nun Study population (n=678)	Analytic sample (n=157)	Excluded participants for missing data; (n=400)[1]	Excluded participants who did not meet criteria for outcomes; (n=121)[2]
Occupation[5]				
Teacher	89.5	93.0	86.9	93.4
House Sister	8.0	6.4	9.8	4.1
Nurse's Aide or Other	2.5	0.6	3.3	2.5
Number of years as a teacher (mean)[6]	37.6	39.8	36.1*	39.6
Linguistic ability[7]				
Quartile of grammatical complexity (%)				
1 (low)	24.4	27.1	20.8	32.3
2	25.0	31.2	24.7	16.1
3	25.6	18.8	29.7	22.6
4 (high)	25.0	22.9	24.7	29.0
Quartile of idea density (%)				
1 (low)	25.0	25.0	23.8	29.0
2	25.6	20.8	28.7	22.6
3	23.9	27.1	23.8	19.3
4 (high)	25.6	27.1	23.8	29.0

*p <0 .05; **p<0.01; ***p<0.001

[1] Tests of significance with this group were conducted in comparison to the analytic sample. Participants in this group include those missing information on key variables in addition to those excluded for not meeting the outcome criteria.

[2] Tests of significance with this group were conducted in comparison to the analytic sample. Participants in this group were any that had complete data but did not meet criteria for the outcomes of this investigation.

[3] In the total Nun Study population, n=507 had complete data concerning the # of languages spoken.

[4] In the total Nun Study population, n=619 had complete data concerning ApoE-E4 status.

[5] In the total Nun Study population, n=676 had complete data concerning occupation.

[6] With respect to participants in the total Nun Study population who were teachers (n=605).

[7] With respect to participants in the total Nun Study population who had handwritten autobiographies analysed for linguistic ability (n=180).

Table 3. Comparison of sub-sample used in the linguistic ability sensitivity analysis and excluded participants (Research Question 1).

Variable	Total Nun Study population (n=678)	Excluded participants[1] (n=521)	Participants without linguistic ability data [2] (n=111)	Analytic sample for sensitivity analysis[3] (n=46)
# of languages spoken[4] (%)				
1	29.6	30.0	28.8	28.3
2	50.5	50.6	52.2	45.6
3	15.6	14.9	14.4	23.9
4	2.4	2.3	1.8	0.0
5	2.0	1.7	2.7	2.2
Educational category (%)				
Grade school	10.0	10.9	9.9	0.0
High school	5.5	6.1	2.7	4.3
Bachelor's degree	39.8	39.6	39.6	43.5
Master's degree+	44.7	43.4	47.7	52.2

*p <0 .05; **p<0.01; ***p<0.001

[1] This group included participants excluded from the original analytic sample; therefore these participants were also excluded from the sensitivity analysis examining linguistic ability from the analytic sample. Comparisons were drawn between this group and the sub-sample of participants with linguistic ability information.

[2] Participants included in the original analytic sample, but were excluded from the linguistic ability sub-sample as they did not have complete information on the linguistic ability variables, or they were removed from the sub-sample due to being influential outliers (n=2). Comparisons were between this group and the linguistic ability sub-sample.

[3] This sample included all participants in the original analytic sample who had complete information concerning linguistic ability.

[4] In the total population, n=507 had complete data concerning the # of languages spoken.

[5] In the total population, n=619 had complete data concerning ApoE-E4 status.

[6] In the total population, n=676 had complete data concerning occupation. Differences across this variable were compared between sub-samples derived from original analytic sample only.

[7] Only participants with teaching occupations (n=605) contributed to this variable in the total population. Differences across this variable were not compared.

[8] The differences across this variable were compared only between the linguistic ability sub-sample and those who were excluded for missing data on linguistic ability variables.

Table 3 (continued). Comparison of sub-sample used in the linguistic ability sensitivity analysis and excluded participants (Research Question 1).

Variable	Total Nun Study population (n=678)	Excluded participants[1] (n=521)	Participants without linguistic ability data[2] (n=111)	Analytic sample for sensitivity analysis[3] (n=46)
Immigrant to US (%)	6.1	6.5	6.3	0.0
% with an ApoE-E4 allele[5]	22.8	21.7	25.2	28.3
Age at death (mean)	90.4	90.2***	92.0***	88.4
Age at last cognitive assessment (mean)	89.5	89.3*	91.2***	87.7
Occupation[6]				
Teacher	89.5	88.4	90.1*	100.0
House Sister	8.0	8.5	9.0*	0.0
Nurse's Aide or Other	2.5	3.1	0.9*	0.0
# of years teaching (mean)[7]	37.6	36.9**	39.1	41.7
% with an outcome of Alzheimer disease[8]	-	-	47.7*	26.1

*p <0 .05; **p<0.01; ***p<0.001

[1] This group included participants excluded from the original analytic sample; therefore these participants were also excluded from the sensitivity analysis examining linguistic ability from the analytic sample. Comparisons were drawn between this group and the sub-sample of participants with linguistic ability information.

[2] Participants included in the original analytic sample, but were excluded from the linguistic ability sub-sample as they did not have complete information on the linguistic ability variables, or they were removed from the sub-sample due to being influential outliers (n=2). Comparisons were between this group and the linguistic ability sub-sample.

[3] This sample included all participants in the original analytic sample who had complete information concerning linguistic ability.

[4] In the total population, n=507 had complete data concerning the # of languages spoken.

[5] In the total population, n=619 had complete data concerning ApoE-E4 status.

[6] In the total population, n=676 had complete data concerning occupation. Differences across this variable were compared between sub-samples derived from original analytic sample only.

[7] Only participants with teaching occupations (n=605) contributed to this variable in the total population. Differences across this variable were not compared.

[8] The differences across this variable were compared only between the linguistic ability sub-sample and those who were excluded for missing data on linguistic ability variables.

Table 4. Comparison of sub-sample of teachers to excluded participants (Research Question 1).

Variable	Total Nun Study population (n=678)	Excluded participants[1] (n=521)	Participants who were not teachers[2] (n=11)	Sub-sample of teachers[3] (n=146)
# of languages spoken[4] (%)				
1	29.6	30.0	18.2	29.4
2	50.5	50.6	81.8	47.9
3	15.6	14.9	0.0	18.5
4	2.4	2.3	0.0	1.4
5	2.0	1.7	0.0	2.7
Educational category (%)				
Grade school	10.0	10.9***	90.9***	0.7
High school	5.5	6.1***	0.0***	3.4
Bachelor's degree	39.8	39.6***	0.0***	43.8
Master's degree+	44.7	43.4***	9.1***	52.0

*p <0 .05; **p<0.01; ***p<0.001

[1] This group included participants excluded from the original analytic sample; therefore these participants were also excluded from the sensitivity analysis examining exclusively teachers from the analytic sample. Comparisons were between this group and the sub-sample composed of teachers.

[2] Participants included in the original analytic sample, but were excluded from the sub-sample of teachers since they held different occupations. Comparisons were between this group and the sub-sample composed of teachers.

[3] This sample included all participants in the original analytic sample that held teaching occupations.

[4] In the total population, n=507 had complete data concerning the # of languages spoken.

[5] In the total population, n=619 had complete data concerning ApoE-E4 status.

6 In the total population, n=676 had complete data concerning occupation. Differences across this variable were compared between sub-samples derived from original analytic sample only.

[7] Only participants with teaching occupations contributed to this variable in the total population. Differences across this variable were not compared.

[8] The differences across this variable were compared only between the sub-sample of teachers and the participants excluded from this group due to holding alternate occupations.

Table 4 (continued). Comparison of sub-sample of teachers to excluded participants (Research Question 1).

Variable	Total Nun Study population (n=678)	Excluded participants[1] (n=521)	Participants who were not teachers[2] (n=11)	Sub-sample of teachers[3] (n=46)
Immigrant to US (%)	6.1	6.5	18.2	3.4
% with an ApoE-E4 allele[5]	22.8	21.7	18.2	26.7
Age at death (mean)	90.4	90.2	94.1*	90.7
Age at last cognitive assessment (mean)	89.5	89.3	93.4*	89.9
Occupation[6]				
Teacher	89.5	88.4	-	100.0
House Sister	8.0	8.5	90.9	-
Nurse's Aide or Other	2.5	3.1	9.1	-
# of years teaching (mean)[7]	37.6	36.9	-	42.7
% with an outcome of Alzheimer disease[8]	-	-	54.5	40.4

*p <0 .05; **p<0.01; ***p<0.001

[1] This group included participants excluded from the original analytic sample; therefore these participants were also excluded from the sensitivity analysis examining exclusively teachers from the analytic sample. Comparisons were between this group and the sub-sample composed of teachers.

[2] Participants included in the original analytic sample, but were excluded from the sub-sample of teachers since they held different occupations. Comparisons were between this group and the sub-sample composed of teachers.

[3] This sample included all participants in the original analytic sample that held teaching occupations.

[4] In the total population, n=507 had complete data concerning the # of languages spoken.

[5] In the total population, n=619 had complete data concerning ApoE-E4 status.

6 In the total population, n=676 had complete data concerning occupation. Differences across this variable were compared between sub-samples derived from original analytic sample only.

[7] Only participants with teaching occupations contributed to this variable in the total population. Differences across this variable were not compared.

[8] The differences across this variable were compared only between the sub-sample of teachers and the participants excluded from this group due to holding alternate occupations.

Table 5. Comparison of analytic sample and those excluded from the analytic sample (Research Question 2).

Variable	Total Nun Study population (n=678)	Analytic sample (n=325)	Excluded participants (n=353)[1]
# of languages spoken[2] (%)			
1	29.6	26.8	34.8
2	50.5	52.6	47.0
3	15.6	15.4	16.0
4	2.4	3.1	1.1
5	2.0	2.1	1.1
Educational category (%)			
Grade school	10.0	4.6	14.9***
High school	5.5	4.3	6.6***
Bachelor's degree	39.8	37.8	42.0***
Master's degree+	44.7	53.2	36.6***
Immigrant to US (%)	6.1	5.5	6.3
% with an ApoE-E4 allele[3]	22.8	18.1	28.0**
Age (in years) at baseline cognitive assessment (mean)	90.4	82.4	84.2***
Occupation other than teacher[4]	10.5	5.5	15.0

*p <0 .05; **p<0.01; ***p<0.001

[1] Tests of significance with this group were conducted in comparison to the analytic sample. Participants in this group include those missing information on key variables in addition to those excluded for being demented at baseline, not having information from at least two cognitive assessments, and those displaying "back-transition" behaviour.

[2] In the total Nun Study population, n=507 had complete data concerning the # of languages spoken.

[3] In the total Nun Study population, n=619 had complete data concerning ApoE-E4 status.

[4] In the total Nun Study population, n=676 had complete data concerning occupation. Examples of other occupations include house sister and nurse's aide.

Description of participants who were excluded from the analytic sample (Research Question 2) due to "back-transition" behaviour

There were 40 participants in the total study population (40/678 participants) who displayed "back-transition" behaviour, meaning they transitioned from clinically demented to non-demented at some point during the study. Some participants displaying this behaviour back-transitioned once before death or study completion while others back-transitioned several times before death or study completion. Since these participants could not be definitively classified as to time of dementia onset and the number of back-transitions varied across the group of participants, all participants displaying this behaviour were excluded from the analyses.

Of the 40 participants who displayed this behaviour, 27 eventually reverted back to a diagnosis of dementia and remained with a diagnosis of dementia until study completion or death. The number of cognitive assessments without dementia in between dementia diagnoses varied in this group of participants: 17 participants had dementia diagnoses separated by one cognitive assessment without dementia; 6 participants had dementia diagnoses separated by two cognitive assessments without dementia; 1 participant had dementia diagnoses separated by four cognitive assessments without dementia; and 3 participants went back and forth several times between not having dementia and having cognition consistent with a diagnosis of dementia.

There were 6 participants who reverted back to being classified as without dementia for one assessment and then died before the next consecutive cognitive assessment. There were 7 participants who reverted back to not having dementia, and remained without dementia, for multiple cognitive assessments until their eventual death.

Table 6. Comparison of sub-sample used in the linguistic ability sensitivity analysis and excluded participants (Research Question 2).

Variable	Total Nun Study population (n=678)	Sensitivity analysis sub-sample participants[1] (n=40)	Participants without linguistic ability data[2] (n=157)	All excluded participants[3] (n=638)
# of languages spoken[4] (%)				
1	29.6	31.2	26.0	29.2
2	50.5	42.4	55.7	52.7
3	15.6	21.7	14.0	14.0
4	2.4	2.8	2.6	2.3
5	2.0	1.9	1.7	1.8
Educational category (%)				
Grade school	10.0	0.0	11.7***	11.8***
High school	5.5	3.8	6.0***	5.8***
Bachelor's degree	39.8	36.8	40.3***	40.4***
Master's degree+	44.7	59.4	41.9***	42.0***
Immigrant to US (%)	6.1	0.0	8.3**	6.1**
% with an ApoE-E4 allele[5]	22.8	19.8	21.4	23.4
Age at baseline cognitive assessment (mean)	90.4	79.8	84.4***	83.9***
Occupation other than teacher[6] (%)	10.5	1.9	12.1***	12.0***

*p <0 .05; **p<0.01; ***p<0.001

[1] This group included participants in the Research Question 2 analytic sample; these participants also had full linguistic ability data.

[2] Participants in this category met inclusion for the original Research Question 2 analytic sample but did not have full data on linguistic ability. Comparisons were drawn between the participants in this group and the participants with linguistic ability data.

[3] Participants in this category were not included in the original Research Question 2 analytic sample (i.e., were missing data on important variables, had dementia at baseline, or were missing data from at least two cognitive assessments). Comparisons were drawn between the participants in this group and participants with linguistic ability data.

[4] In the total population, n=507 had complete data concerning the # of languages spoken.

[5] In the total population, n=619 had complete data concerning ApoE-E4 status.

[6] In the total population, n=676 had complete data concerning occupation. Examples of other occupations include house sister and nurse's aide.

Appendix D: Languages spoken by Nun Study participants.

Table 1. Languages spoken by the total Nun Study population, and by participants included in the different analytic samples.

Language	% of total population with data (n=507)	% of analytic sample (Research Question 1; n=157)	% of analytic sample (Research Question 2; n=325)
English	100.0	100.0	100.0
German	41.4	38.8	40.9
French	18.3	22.9	22.1
Spanish	12.6	14.6	12.3
Polish	10.8	6.4	12.6
Italian	3.2	5.7	4.0
Latin	4.9	4.5	4.6
Slovak	1.8	3.2	1.5
Czech	1.8	1.9	1.2
Japanese	0.4	0.0	0.3
Chamorro	0.4	0.0	0.3
Other	1.0	0.6	1.2

Appendix E: Ethics documentation.

UNIVERSITY OF WATERLOO

OFFICE OF RESEARCH ETHICS

Notification of Ethics Clearance of Application to Conduct Research with Human Participants

---

Principal/Co-Investigator: Suzanne Tyas
Department: Health Studies & Gerontology

Student Investigator: Erica Hack
Department: Health Studies & Gerontology

ORE File #: 16551

Project Title: Early and late-life predictors of symptomatic and presymptomatic Alzheimer's disease in the Nun Study

---

This certificate provides confirmation that the additional information/revised materials requested for the above project have been reviewed and are considered acceptable in accordance with the University of Waterloo's Guidelines for Research with Human Participants and the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans. Thus, the project now has received ethics clearance. This clearance is valid for a period of ***four years*** *from the date shown below and is subject to an* ***annual ethics review process*** *(see Note 2). A new application must be submitted for on-going projects continuing beyond four years.*

Note 1*: This project must be conducted in accordance with the description in the application and revised materials for which ethics clearance has been granted. All subsequent modifications to the application must be submitted for prior ethics review using ORE Form 104 and must not be intiated until notification of ethics clearance has been received.*

Note 2*: All ongoing research projects must undergo annual ethics review. ORE Form 105 is used for this purpose and must be submitted by the Faculty Investigator/Supervisor (FI/FS) when requested by the ORE. Researchers must submit a Form 105 at the conclusion of the project if it continues for less than a year.*

Note 3*: FIs and FSs also are reminded that they must immediately report to the ORE (using ORE Form 106) any events related to the procedures used that adversely affected the participants and the steps taken to deal with these.*

Susan E. Sykes, Ph.D., C.Psych.
Director, Office of Research Ethics

OR
Susanne Santi, M. Math
Senior Manager, Research Ethics

OR
Julie Joza, B.Sc.
Manager, Research Ethics

7/16/2010
Date

Appendix F. Description of discrete-time units utilized by Research Question 2.

Table 1. Time elapsed between each cognitive assessment, for each participant included in the analytic sample (Research Question 2).

	Average length of time between cognitive assessment per participant (in years; n=325)
Mean	1.51
Standard Deviation	0.32
Median	1.46
Mode	1.47
Range	3.83
Inter-quartile Range	0.18

Appendix G: Research Question 1 logistic regression models.

Table 1. Logistic regression models employed by the Research Question 1 analyses.

Analysis	Model	
Main analyses	**Model 1a:**	AD = Multilingualism (four-level definition[1])
	Model 1b:	AD = Multilingualism (four-level definition[1]) + Age at last cognitive assessment + educational level + immigrant status + ApoE-E4 status
	Model 2a:	AD = Multilingualism (two-level definition[2])
	Model 2b:	AD = Multilingualism (two-level definition[2]) + Age at last cognitive assessment + educational level + immigrant status + ApoE-E4 status
Sensitivity analysis 1: linguistic ability	**Model 3:**	AD = Multilingualism (three-level definition[3]) + ApoE-E4 status, age at last cognitive assessment
	Model 4:	AD = Multilingualism (two-level definition[2]) + Quartile of grammatical complexity
Sensitivity analysis 2: restricted to teachers	**Model 5:**	AD = Multilingualism (four-level definition[1]) + Age at last cognitive assessment + level of attained formal education + teaching career length + immigrant status + ApoE-E4 status
	Model 6:	AD = Multilingualism (two-level definition[2]) + Age at last cognitive assessment + level of attained formal education + teaching career length + immigrant status + ApoE-E4 status

[1] Participants were classified as speaking one language (reference group), two languages, three languages, or four or more languages.
[2] Participants were classified as either speaking two or more languages or speaking one language (reference group).
[3] Participants were classified as speaking one language (reference group), two languages, or three or more languages.

Appendix H: Research Question 1 linguistic ability sensitivity analysis backward elimination summaries, odds ratio estimates, and descriptions of influential outliers omitted from the analyses.

Table 1. Results of backward elimination procedure for the linguistic ability sensitivity analysis: three-level multilingualism variable.

Effect retained by model	**p-value**
Multilingualism	0.10
Age at last cognitive assessment	0.10
ApoE-E4 status	0.0006
Effect removed from model	**p-value**
Level of attained formal education	0.98
# of years as a teacher	1.00
Grammatical complexity	0.68
Idea density	0.98

Variables were removed by the procedure if they did not meet a significance level of 0.15.
Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele

Table 2. Results of backward elimination procedure for the linguistic ability sensitivity analysis: two-level multilingualism variable (before omission of any outliers).

Effect retained by model	**p-value**
Multilingualism	0.06
Grammatical complexity	0.14
ApoE-E4 status	0.001
Effect removed from model	**p-value**
Level of attained formal education	0.69
# of years as a teacher	0.58
Age at last cognitive assessment	0.16
Idea density	0.55

Variables were removed by the procedure if they did not meet a significance level of 0.15.
Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele

Description of outliers and their influence on the results

When all participants from the Research Question 1 analytic sample with complete linguistic ability information (n=48) were analysed using the backward elimination procedure, three variables were judged to be significant and were thus retained by the logistic regression model predicting AD development. These variables were multilingualism, ApoE-E4 status, and grammatical complexity. Speaking two or more languages was associated with a significant decrease in AD risk, compared to speaking one language while the highest quartile of grammatical complexity was associated with a significant decrease in AD risk compared to the lowest quartile of grammatical complexity (Table 4 in Appendix H). The OR estimate associated with possessing an ApoE-E4 allele was not computable due to quasi-separation of the data. This model also showed significantly poor fit to the data according to the Hosmer-Lemeshow goodness-of-fit test. One participant was identified to be more influential than the other participants using standard diagnostic residual plots and criteria (see Section 5.2.1), and was subsequently removed from the sample. The excluded participant was a highly educated (graduate degree; highest category of educational attainment), non-immigrant teacher who spoke two languages fluently. She was 76.8 years old at baseline, developed clinical dementia at her seventh follow-up assessment, and was 88.5 years old at her last cognitive assessment (eighth follow-up visit). She died at 89.1 years of age. This participant did not possess any ApoE-E4 alleles and had neuropathology characteristic of AD. This participant ranked in the second and third quartiles with regard to idea density and grammatical complexity, respectively. This participant may have been an outlier since she was highly educated, had high grammatical complexity, and did not have an ApoE-E4 allele—all generally associated with a reduced risk of AD—yet still developed AD.

After the removal of this participant, the backward elimination procedure was again performed. The same variables (multilingualism, ApoE-E4 status, and grammatical complexity) remained significant predictors in the model. Speaking two or more languages was, again, significantly associated with a decrease in AD risk, compared to speaking one language. The OR estimates for both ApoE-E4 status and grammatical complexity, however, were both not calculated due to quasi-complete separation. This model also had a small (but not statistically significant) Hosmer-Lemeshow goodness-of-fit test statistic, suggesting a poor fit with the data. After examining the residual diagnostics once more, another disproportionately influential outlier was identified and removed from the sample.

The second participant excluded from the sample was a highly educated (graduate degree; highest category of educational attainment), non-immigrant who was not a teacher and spoke only one language fluently. She was 83.3 years old at baseline, remained cognitively intact, and was 88.9 years of age at her last cognitive assessment (fourth follow-up visit). This participant died at 90.1 years of age, and did not possess any ApoE-E4 alleles. This participant ranked in the third and first quartiles with regard to idea density and grammatical complexity, respectively. This participant may have been an outlier since she held an occupation other than a teacher (the only one to do so in this particular sample). Another potential explanation is that she had a low grammatical complexity score, yet remained dementia-free. Since a low grammatical complexity score was generally associated with increased AD risk, this quality may have led this participant to be an outlier from the rest of the sample.

After the exclusion of these two participants, ApoE-E4 status was no longer a significant predictor of AD development. The association between multilingualism and AD weakened and was no longer statistically significant. High grammatical complexity, on the other hand,

generally remained significantly related to a decrease in AD risk once the outliers were removed from the sample.

Table 3. Results of backward elimination procedure for the linguistic ability sensitivity analysis: two-level multilingualism variable (after omission of two outliers).

Effect retained by model	**p-value**
Multilingualism	0.10
Grammatical complexity	0.14
Effect removed from model	
ApoE-E4 status	0.91
Level of attained formal education	0.95
# of years as a teacher	0.89
Age at last cognitive assessment	0.93
Idea density	1.00

Variables were removed by the procedure if they did not meet a significance level of 0.15.
Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele

Appendix I: Discrete-time survival analysis models, estimates of dementia hazard probabilities, and dementia odds ratios used in addressing Research Question 2.

Table 1. Logistic regression models employed by the discrete-time survival analyses in Research Question 2.

Analysis	Model	
Main discrete-time survival analyses (with time variables)	**Model 7:**	Dementia = Multilingualism (four-level definition[1]) + Transition period 1-11[2]
	Model 8:	Dementia = ApoE-E4 status + Transition period 1-11[2]
	Model 9:	Dementia = Age at baseline cognitive assessment (3-level categorical variable[3]) + Transition period 1-11[2]
	Model 10:	Dementia = Multilingualism (four-level definition[1]) + ApoE-E4 status + Age at baseline cognitive assessment (3-level categorical variable[3]) + Transition period 1-11[2]
Main discrete-time survival analyses (models derived from log likelihood test)	**Model 11:**	Dementia = Multilingualism (four-level definition[1]) + ApoE-E4 status + Age at baseline cognitive assessment (3-level categorical variable[3])
	Model 12:	Dementia = Multilingualism (four-level definition[1]) + ApoE-E4 status + Age at baseline cognitive assessment (continuous variable)
Sensitivity analysis: linguistic ability	**Model 13:**	Dementia = Multilingualism (four-level definition[1]) + ApoE-E4 status + educational level (three-level definition[4]) + Quartile of idea density
	Model 14:	Dementia = Multilingualism (four-level definition[1]) + ApoE-E4 status+ educational level (three-level definition[4]) + Quartile of idea density (two-level definition[6])

[1] Participants were classified as speaking one language (reference group), two languages, three languages, or four or more languages.

[2] Where a transition period was defined as the period of time between consecutive cognitive assessments; there were 11 transition periods assessed (as there were 12 cognitive assessments in the Nun Study).

[3] Participants were divided into three categories based on their ages at baseline cognitive assessment: the first age category included participants aged younger than 80 years at baseline; the second age category included participants aged between 80 and less than 85 years at baseline; and the third age category included participants aged 85+ years at baseline.

[4] Participants were divided into three educational categories; the reference group consisted of participants with a high school level education or less: the second group consisted of

participants with a Bachelor`s degree; and the third group consisted of participants with a Master's degree or higher.

[5] Participants were divided into three categories based on their ranked score for idea density: the first category corresponded to participants scoring in the first quartile (reference group); the second category corresponded to participants scoring in the second quartile; and the third category corresponded to participants scoring in either of the top two quartiles.

[6] Participants were divided into two categories based on their ranked score for idea density: the first category corresponded to participants scoring in the first quartile (reference group); the second category corresponded to participants scoring in the top three quartiles (second, third, or fourth quartiles).

Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele

Table 2. Dementia hazard probability estimates for each category of multilingualism, generated by Model 7.

Transition period	Hazard estimate			
	Speaking one language	**Speaking two languages**	**Speaking three languages**	**Speaking four or more languages**
1	0.057330071	0.089696	0.064957	0.009412375
2	0.040913303	0.064011	0.046356	0.006717092
3	0.070948572	0.111003	0.080387	0.011648242
4	0.055994601	0.087607	0.063444	0.009193119
5	0.069481135	0.108707	0.078725	0.01140732
6	0.054063284	0.084585	0.061256	0.008876038
7	0.033399979	0.052256	0.037843	0.005483564
8	0.050671077	0.079278	0.057412	0.00831911
9	0.04648619	0.07273	0.05267	0.007632041
10	0.060889167	0.095264	0.06899	0.009996702
11	0.057160053	0.08943	0.064764	0.009384461

Table 3. Dementia hazard odds ratios generated from Model 7.

Parameter	**OR**	**95% CI**
Multilingualism		
Speaking two languages vs. one	1.56	1.00, 2.52
Speaking three languages vs. one	1.13	0.58, 2.14
Speaking four or more languages vs. one	**0.16**	**0.01, 0.78**
Transition period[1]		
1	1.00	0.33, 4.38
2	0.72	0.22, 3.20
3	1.24	0.40, 5.46
4	0.98	0.30, 4.42
5	1.22	0.37, 5.49
6	0.95	0.26, 4.48
7	0.58	0.12, 3.08
8	0.89	0.21, 4.50
9	0.81	0.17, 4.31
10	1.07	0.22, 5.66

Bolded estimates indicate statistical significance.

[1] No estimate was available for transition period 11 as the parameter estimate was equal to zero.

Abbreviations: OR = odds ratio estimate; 95% CI = 95% confidence interval

Table 4. Dementia hazard probability estimates generated by Model 8.

Transition period	Hazard estimate	
	ApoE-E4 allele absent	**ApoE-E4 allele(s) present**
1	0.060313459	0.131782822
2	0.042638401	0.093163432
3	0.074593642	0.162984528
4	0.059261229	0.129483736
5	0.074982539	0.163834255
6	0.05895387	0.128812168
7	0.03601211	0.078685216
8	0.054030855	0.11805555
9	0.049040931	0.107152738
10	0.062249797	0.136013654
11	0.063666294	0.13910865

Table 5. Dementia hazard odds ratios generated by Model 8.

Parameter	**OR**	**95% CI**
ApoE-E4 status		
Possession of an ApoE-E4 allele	**2.18**	**1.36, 3.41**
Transition period[1]		
1	0.95	0.31, 4.31
2	0.67	0.21, 3.00
3	1.17	0.38, 5.15
4	0.93	0.28, 4.19
5	1.18	0.36, 5.31
6	0.93	0.25, 4.38
7	0.57	0.12, 3.00
8	0.85	0.20, 4.30
9	0.77	0.16, 4.07
10	0.98	0.20, 5.19

Bolded estimates indicate statistical significance.
[1] No estimate was available for transition period 11 as the parameter estimate was equal to zero.
Abbreviations: OR = odds ratio estimate; 95% CI = 95% confidence interval; ApoE-E4 = Apolipoprotein E-E4 allele

Table 6. Dementia hazard estimates generated by Model 9 (Research Question 2).

Transition period	Hazard estimates		
	Baseline age less than 80 years	**Baseline age between 80 and less than 85 years**	**Baseline age 85 years and above**
1	0.037442832	0.060698275	0.157608684
2	0.028005902	0.045400144	0.117885673
3	0.049725371	0.080609401	0.209309764
4	0.040427703	0.065537027	0.170172951
5	0.052657848	0.085363216	0.221653484
6	0.043851779	0.071087768	0.184585963
7	0.026626721	0.043164365	0.112080266
8	0.039774073	0.064477433	0.167421616
9	0.036376221	0.0589692	0.153118982
10	0.046768752	0.075816449	0.196864422
11	0.049186307	0.07973553	0.207040677

Table 7. Hazard odds ratios generated by Model 9 (Research Question 2).

Parameter	**OR**	**95% CI**
Age category[1]		
Age between 80 and less than 85 years at baseline	1.62	0.99, 2.67
Age 85+ at baseline	**4.21**	**2.57, 6.97**
Transition period[2]		
1	0.76	0.24, 3.35
2	0.57	0.17, 2.57
3	1.01	0.32, 4.49
4	0.82	0.25, 3.74
5	1.07	0.32, 4.88
6	0.89	0.24, 4.25
7	0.54	0.11, 2.88
8	0.81	0.19, 4.14
9	0.74	0.15, 3.95
10	0.95	0.20, 5.08

Bolded estimates indicate statistical significance.
[1] Where the reference category included participants aged less than 80 years at baseline.
[2] No estimate was available for transition period 11 as the parameter estimate was equal to zero.
Abbreviations: OR = odds ratio estimate; 95% CI = 95% confidence interval

Table 8. Dementia Hazard probability estimates generated by Model 10 (Research Question 2).

Hazard Probability Estimates						
Time	**ApoE4[1]=0 Langs.[2]=1 Age[3]=1**	**ApoE4=0 Langs.=1 Age=2**	**ApoE4=0 Langs.=1 Age=3**	**ApoE4=0 Langs.=2 Age=1**	**ApoE4=0 Langs.=2 Age=2**	**ApoE4=0 Langs.=2 Age=3**
1	0.022	0.036	0.105	0.034	-2.871	0.163
2	0.017	0.028	0.080	0.026	-3.146	0.124
3	0.031	0.051	0.148	0.048	-2.525	0.231
4	0.026	0.043	0.125	0.041	-2.693	0.195
5	0.035	0.058	0.166	0.054	-2.410	0.259
6	0.030	0.049	0.142	0.046	-2.564	0.222
7	0.018	0.031	0.088	0.029	-3.041	0.138
8	0.028	0.047	0.135	0.044	-2.614	0.211
9	0.026	0.043	0.125	0.041	-2.690	0.196
10	0.035	0.058	0.169	0.055	-2.395	0.263
Time	**ApoE4=0 Langs.=3 Age=1**	**ApoE4=0 Langs.=3 Age=2**	**ApoE4=0 Langs.=3 Age=3**	**ApoE4=0 Langs.=4 Age=1**	**ApoE4=0 Langs.=4 Age=2**	**ApoE4=0 Langs.=4 Age=3**
1	0.027	0.045	0.130	0.003	0.005	-4.267
2	0.021	0.034	0.099	0.002	0.004	-4.541
3	0.038	0.064	0.184	0.004	0.007	-3.920
4	0.032	0.054	0.156	0.003	0.006	-4.089
5	0.043	0.072	0.207	0.005	0.008	-3.806
6	0.037	0.061	0.177	0.004	0.007	-3.960
7	0.023	0.038	0.110	0.002	0.004	-4.436
8	0.035	0.058	0.169	0.004	0.006	-4.009
9	0.033	0.054	0.156	0.003	0.006	-4.086
10	0.044	0.073	0.210	0.005	0.008	-3.790
11	0.027	0.045	0.130	0.003	0.005	-4.267
Time	**ApoE4=1 Langs.=1 Age=1**	**ApoE4=1 Langs.=1 Age=2**	**ApoE4=1 Langs.=1 Age=3**	**ApoE4=1 Langs.=2 Age=1**	**ApoE4=1 Langs.=2 Age=2**	**ApoE4=1 Langs.=2 Age=3**
1	0.055	0.092	0.265	0.086	0.143	0.414
2	0.042	0.070	0.201	0.065	0.109	0.314
3	0.078	0.130	0.375	0.122	0.203	0.585
4	0.066	0.110	0.317	0.103	0.171	0.494
5	0.087	0.146	0.420	0.137	0.227	0.656
6	0.075	0.125	0.360	0.117	0.195	0.562
7	0.047	0.078	0.224	0.073	0.121	0.349
8	0.071	0.119	0.343	0.111	0.185	0.535
9	0.066	0.110	0.318	0.103	0.172	0.496
10	0.089	0.148	0.427	0.139	0.231	0.666
11	0.094	0.157	0.453	0.147	0.245	0.707

Table 8. Dementia Hazard probability estimates generated by Model 10 (Research Question 2).

Hazard Probability Estimates						
Time	**ApoE4=1 Langs.=3 Age=1**	**ApoE4=1 Langs.=3 Age=2**	**ApoE4=1 Langs.=3 Age=3**	**ApoE4=1 Langs.=4 Age=1**	**ApoE4=1 Langs.=4 Age=2**	**ApoE4=1 Langs.=4 Age=3**
1	0.069	0.114	0.330	0.007	0.012	-3.338
2	0.052	0.087	0.251	0.006	0.009	-3.612
3	0.097	0.162	0.467	0.010	0.017	-2.991
4	0.082	0.137	0.394	0.009	0.015	-3.160
5	0.109	0.181	0.523	0.012	0.020	-2.877
6	0.093	0.155	0.448	0.010	0.017	-3.031
7	0.058	0.097	0.279	0.006	0.010	-3.507
8	0.089	0.148	0.427	0.010	0.016	-3.080
9	0.082	0.137	0.395	0.009	0.015	-3.157
10	0.111	0.184	0.531	0.012	0.020	-2.861
11	0.117	0.195	0.564	0.013	0.021	-2.802

[1]ApoE4=0 refers to the absence of ApoE-E4 alleles; ApoE4=1 refers to the presence of an ApoE-E4 allele.
[2]Langs. refers to the number of languages spoken.
[3]Age refers to the age category, where age category 1 was composed of participants younger than 80 years at baseline assessment; age category 2 refers to participants aged between 80 and younger than 85 years at baseline; and age category 3 refers to participants aged more than 85+ years at baseline assessment.

Table 9. Hazard probability estimates generated by Model 11 (Research Question 2).

	ApoE-E4 allele absent			
	Baseline age less than 80 years			
	1 language	*2 languages*	*3 languages*	*4+ languages*
Hazard probability estimate	0.027	0.041	0.033	0.004
	Baseline age between 80 and less than 85 years			
	1 language	*2 languages*	*3 languages*	*4+ languages*
Hazard probability estimate	0.043	0.066	0.053	0.006
	Baseline age 85+ years			
	1 language	*2 languages*	*3 languages*	*4+ languages*
Hazard probability estimate	0.119	0.182	0.146	0.016
	ApoE-E4 allele(s) present			
	Baseline age less than 80 years			
	1 language	*2 languages*	*3 languages*	*4+ languages*
Hazard probability estimate	0.064	0.098	.078	0.009
	Baseline age between 80 and less than 85 years			
	1 language	*2 languages*	*3 languages*	*4+ languages*
Hazard probability estimate	0.104	0.160	0.127	0.014
	Baseline age 85+ years			
	1 language	*2 languages*	*3 languages*	*4+ languages*
Hazard probability estimate	0.286	0.438	0.350	0.039

Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele

Table 10. Hazard odds ratio estimates generated by Model 11 (Research Question 2).

Parameter	OR	95% CI
Multilingualism		
Speaking two languages vs. one	1.53	0.96, 2.50
Speaking three languages vs. one	1.22	0.62, 2.35
Speaking four or more languages vs. one	**0.14**	**0.01, 0.66**
ApoE-E4 status		
E4 allele(s) present vs. absent	**2.40**	**1.48, 3.81**
Age[1]		
Between 80 and less than 85 years at baseline	**1.63**	**1.00, 2.69**
85+ years at baseline	**4.48**	**2.74, 7.39**

Bolded values indicate statistical significance.
[1] Where the reference category consisted of participants aged less than 80 years at baseline.
Abbreviations: OR = odds ratio estimate; 95% CI = 95% confidence interval; ApoE-E4 = Apolipoprotein E-E4 allele

Table 11. Dementia hazard odds ratio estimates generated by Model 12 (Research Question 2).

Parameter	OR	95% CI
Multilingualism		
Speaking two languages vs. one	1.15	0.95, 2.47
Speaking three languages vs. one	1.29	0.65, 2.47
Speaking four or more languages vs. one	0.14	0.01, 0.66
Presence of an ApoE-E4 allele	**2.26**	**1.40, 3.56**
Age at baseline (per year increase)	**1.12**	**1.08, 1.17**

Bolded values indicate statistical significance.
Abbreviations: OR = odds ratio estimate; 95% CI = 95% confidence interval; ApoE-E4 = Apolipoprotein E-E4 allele

Table 12. Dementia hazard odds ratio estimates generated by Model 13 (Research Question 2).

Parameter	**OR**	**95% CI**
Multilingualism		
Speaking two languages vs. one	1.21	0.46, 3.14
Speaking three languages vs. one	0.68	0.14, 2.40
Speaking four or more languages vs. one	0.58	0.03, 3.84
Presence of an ApoE-E4 allele	**2.73**	**1.03, 7.18**
Educational level		
Bachelor's degree vs. High school	**0.20**	**0.05, 0.85**
Master's degree+ vs. High school	**0.13**	**0.03, 0.56**
Quartile of idea density[1]		
2 vs. 1	**0.23**	**0.07, 0.72**
3 vs. 1	**0.14**	**0.04, 0.44**
4 vs. 1	**0.15**	**0.04, 0.48**

[1]Where quartile one was the lowest quartile of idea density, and quartile four was the highest.
Bolded values indicate statistical significance.
Abbreviations: OR = odds ratio estimate; 95% CI = 95% confidence interval; ApoE-E4 = Apolipoprotein E-E4 allele

Table 13. Hazard probability estimates generated by Model 14 (Research Question 2).

Hazard probability estimate	**ApoE-E4 allele absent**			
	High school			
	1 language	*2 languages*	*3 languages*	*4+ languages*
Low idea density[1]	0.995	1.075	0.622	0.530
High idea density[2]	0.169	0.183	0.106	0.090
	Bachelor's degree			
	1 language	*2 languages*	*3 languages*	*4+ languages*
Low idea density	0.178	0.192	0.111	0.095
High idea density	0.030	0.033	0.019	0.016
	Master's degree+			
	1 language	**2 languages**	**3 languages**	**4+ languages**
Low idea density	0.120	0.129	0.075	0.064
High idea density	0.020	0.022	0.013	0.011
	ApoE-E4 allele(s) present			
	High school			
	1 language	**2 languages**	**3 languages**	**4+ languages**
Low idea density	2.34	2.52	1.46	1.25
High idea density	0.398	0.430	0.249	0.212
	Bachelor's degree			
	1 language	**2 languages**	**3 languages**	**4+ languages**
Low idea density	0.418	0.451	0.261	0.222
High idea density	0.071	0.077	0.044	0.038
	Master's degree+			
	1 language	**2 languages**	**3 languages**	**4+ languages**
Low idea density	0.281	0.030	0.176	0.150
High idea density	0.048	0.052	0.030	0.026

[1]Where low idea density represents idea density scores from the first quartile.

[2] Where high idea density represents idea density scores from the second, third, and fourth quartiles.

Abbreviations: ApoE-E4 = Apolipoprotein E-E4 allele

Table 14. Dementia hazard odds ratio estimates generated by Model 14 (Research Question 2).

Parameter	**OR**	**95% CI**
Multilingualism		
Speaking two languages vs. one	1.08	0.43, 2.69
Speaking three languages vs. one	0.62	0.16, 2.41
Speaking four languages or more vs. one	0.53	0.06, 4.91
Possession of an ApoE-E4 allele	2.35	0.98, 5.65
Educational level		
Bachelor's degree vs. High school	**0.18**	**0.05, 0.69**
Master's degree+ vs. High school	**0.12**	**0.03, 0.47**
Idea density[1]		
High vs. low	**0.17**	**0.07, 0.42**

[1]Where low idea density represents idea density scores from the first quartile, and where high idea density represents idea density scores from the second, third, and fourth quartiles.
Bolded values indicate statistical significance.
Abbreviations: OR = odds ratio estimate; 95% CI = 95% confidence interval; ApoE-E4 = Apolipoprotein E-E4 allele