The Pennsylvania State University

The Graduate School

College of the Liberal Arts

THE ROLE OF WHITE MATTER INTEGRITY IN AGE-RELATED LANGUAGE PRODUCTION DIFFERENCES

A Thesis in

Psychology

by

Sara B. Winter

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science

May 2018

The thesis of Sara Winter was reviewed and approved* by the following:

Michele T. Diaz
Associate Professor of Psychology
Thesis Advisor

Nancy A. Dennis
Associate Professor of Psychology

Lesley A. Ross
Associate Professor of Human Development and Family Studies

Melvin Mark
Professor of Psychology
Head of the Department of Psychology

*Signatures are on file in the Graduate School

Abstract

Despite having equal comprehension ability, older adults have more language production difficulties than younger adults (Diaz, Johnson, Burke, & Madden, 2014). According to the Transmission Deficit Hypothesis, language difficulties stem from signal transmission failures which increase with age. The hypothesis holds that the one-to-one mapping of the phonological system creates vulnerability to transmission failures but the many-to-one mapping of semantic networks provides protection from effects of transmission failure (Burke and MacKay, 1991). Alternatively, the Inhibition Deficit Hypothesis would posit that age-related declines in inhibition increase the task-demands of speaking, leading to poorer performance (Hasher & Zacks, 1988). Since white matter integrity has been shown to mediate age-behavior relationships, a potential mechanism underlying both accounts may be age-related white matter integrity declines (Head et al, 2004; Bennet & Madden, 2014). This study explored the relationship between white matter integrity and age-related language deficits using Diffusion Tensor Imaging (DTI) to test hypotheses generated by the Transmission Deficit Hypothesis and the Inhibition Deficit Hypothesis. Findings supported the Transmission Deficit Hypothesis; white matter integrity declined across the brain but the relationship between white matter integrity and outcomes only manifest in phonological behaviors and phonological-task activation. Importantly, age mediated the relationships between white matter integrity and behavioral and activation outcomes, suggesting that white matter integrity decline is a substrate of age-related language production deficits.

Table of Contents

List of Figures

List of Tables

Acknowledgements

First and foremost it is necessary to give credit to the many individuals who made this project possible. This work was supported in part by T32 grant AG049676 and by R01 grant AG034138 to Michele Diaz, Ph.D. I would also like to thank the staff and scientists at the Center for Language Science and the Social, Life, and Engineering Sciences Imaging Center where the project was conducted; Dr. Avery Rizio who cleaned the data, performed the tractography on the FAT and provided advice on methods, pipelines, and theory throughout; and Jack Dempsey and Annie Ryder who resiliently helped to examine tract quality. Further, I am grateful to the members of my committee; Michele Diaz, Ph.D, Dr. Nancy Dennis, and Dr. Lesley Ross. Their selfless mentorship has not only made me a better scientist but has also elevated the quality of this project immensely. I am also personally indebted to Priscila L.B., Abbie C., and the Language and Aging lab, Hossein, Haoyun, Victoria, and Anna, whose constant encouragement and friendship has sustained me. Most of all, I give thanks to God, my fiancé, my sister, and my parents, all of whom have loved and supported me every step of the way.

Chapter 1

Introduction

Despite maintaining the ability to comprehend language, older adults experience more language production difficulties than younger adults (Diaz et al., 2014). Two prominent behavioral theories account for these changes in language ability. First, the Transmission Deficit Hypothesis suggests that language production difficulties stem from signal transmission failures which increase with age. However, behavioral manifestations of these failures vary as a result of how knowledge is instantiated in the brain (Burke, MacKay, Worthley, & Wade, 1991). Alternatively, the Inhibition Deficit Hypothesis suggests that, with age, individuals become less able to inhibit irrelevant stimuli. In the case of language, this means that selecting a word from among competitors is more difficult and speech becomes slower and less accurate (Hasher & Zacks, 1988). In addition to providing theoretical explanations of age-related differences in language ability, these theories are supported by a wealth of behavioral evidence.

On the other hand, there is a dearth of exploration on the neural substrates of both the Transmission and Inhibition Deficit Hypotheses in the context of language production. The literature suggests white matter integrity, which declines in healthy aging (Bennett & Madden, 2014; Head et al., 2004) could be a potential source of age-related cognitive deficits in processing speed and executive functions (Bennett & Madden, 2014; Fjell et al., 2017; Gazes et al., 2016; Hedden et al., 2016). However, only a handful of studies have looked at the implications of white matter integrity in older adult’s language ability (Rizio, Moyer, & Diaz, 2017; Stamatakis, Shafto, Williams, Tam, & Tyler, 2011). Thus, this study utilized a Picture Word Interference (PWI) paradigm and Diffusion Tensor Imaging (DTI) to examine the possibility that white matter integrity drives age-related transmission and inhibition deficits and ultimately age-related language production deficits.

Transmission Deficit Hypothesis

In accord with traditional Hebbian brain dynamics and the interactive activation model of language production (Dell & O'Seaghdha, 1992), the Transmission Deficit Hypothesis explains age-related language changes according to the number and strength of connections in the brain. Specifically, the Transmission Deficit Hypothesis claims that increased word finding difficulties in older adults result from a weakening of connections across the entire brain. As connections weaken, signals fail to reach their destination and actions—such as selecting a word—cannot be completed. However, the hypothesis also suggests that behavioral consequences of these transmission failures are not observed in semantic tasks because of the knowledge structure of the semantic system. The semantic system is more interconnected than the phonological system and thus robust to signal transmission failures. If one signal transmission does fail in the semantic system, the signal can still be conveyed via other connections (MacKay & Burke, 1990).

The Transmission Deficit Hypothesis accounts for a myriad of behavioral findings regarding language production ability across the lifespan. For instance, all individuals, including older adults, are faster and more accurate when naming items with a higher word frequency (Rizio et al., 2017). This is concordant with the Transmission Deficit Hypothesis which suggests that if a pathway falls into disuse, connections weaken and information is more difficult to access. Likewise, all individuals are more likely to make proper name errors than errors in naming common nouns such as occupations (James, 2004) since proper names are instantiated in memory as a single, low strength network node. Notably, the rate of proper name errors, but not common name errors, increases in older adults. This too can be explained by the Transmission

Deficit Hypothesis as common nouns are instantiated as nodes with multiple, stronger connections.

Indeed by accounting for both the strength of connections and the number of connections in a network, the Transmission Deficit Hypothesis makes provision for the different trajectories of the phonological and semantic systems in older adulthood. That is, although the Transmission Deficit Hypothesis suggests that connections weaken across the brain, the semantic system has been found to be robust to decline (Diaz et al., 2014). Semantic distractors may even cause more interference in older adults (Taylor & Burke, 2002) and semantic priming may speed lexical decision and pronunciation more with age (Laver & Burke, 1993). In sum, the Transmission Deficit Hypothesis emphasizes that age-related deficits occur across systems of the brain but that the structure of the knowledge system determines behavioral outcomes (Burke & Mackay, 1997).

Inhibition Deficit Hypothesis

In contrast to the emphasis on the knowledge system structure proposed by the Transmission Deficit Hypothesis, the Inhibition Deficit Hypothesis suggests that decline in domain general functions, specifically inhibition, determines behavioral outcomes in aging (Hasher & Zacks, 1988). In the case of language production, the Inhibition Deficit Hypothesis would suggest that word retrieval requires the inhibition of competitors and that older adults are less able to selectively inhibit alternatives than younger adults (Hasher & Zacks, 1988; Zacks & Hasher, 1997). In this way, the Inhibition Deficit Hypothesis predicts that the more inhibition a task requires the more age-related difficulties an individual will encounter. If the Inhibition Deficit Hypothesis holds, stability in the semantic system of older adults would indicate that semantic comprehension does not rely as much on inhibition ability as production does or, even, that semantic comprehension could be facilitated by lack of inhibition.

Though the Inhibition Deficit Hypothesis has not been as extensively studied as the Transmission Deficit Hypothesis in the domain of language, the Inhibition Deficit Hypothesis is supported by some behavioral findings in the language comprehension literature. For instance, across ages, words with high phonological neighborhood density are perceived more slowly than words with few and infrequent phonological neighbors (Chen & Mirman, 2012; Luce & Pisoni, 1998; Vitevitch & Luce, 2016). Words with highly frequent neighbors, which are stronger competitors, are also perceived more slowly than words with less frequent neighbors (Chen & Mirman, 2012; Vitevitch & Luce, 2016). In the domain of production, Britt, Ferrera, and Mirman also provide partial support for the Inhibition Deficit Hypothesis. They found that participants of all ages were slower and less accurate to name words with many semantic competitors (i.e., jam/jelly) but also observed an age by competition interaction, suggesting that naming words with many semantic competitors was especially difficult for older adults compared to younger adults (Britt, Ferrara, & Mirman, 2016). Presumably this is because older adults had difficulty inhibiting semantic competitors.

Conversely, the same study by Britt, Ferrera, and Mirman did not observe significant age differences in the effect of lexical competition (i.e., synonymous concepts like sofa/couch). They only found age differences in semantic selection--selection between similar but distinct concepts (Britt et al., 2016). This runs counter to the prediction that language systems would broadly show decrement as a result of lower inhibition (Hasher & Zacks, 1988). Thus, though the Inhibition Deficit Hypothesis accounts for some observed age-related language changes, it does not account for the stability seen in the semantic system (Burke, 1997) as the interference effect of competition is evident more at the level of semantic selection than lexical selection in older adults. Also counter to the predictions of the Inhibition Deficit Hypothesis are the facilitative effects of phonological neighborhood density on word production (Vitevitch & Luce, 2016). Words with highly dense phonological neighborhoods are named more quickly than word with sparse neighborhoods (Chen & Mirman, 2012; Vitevitch & Luce, 2016). Moreover, the Inhibition Deficit hypothesis predicts that older adults should report more competing alternatives when in a tip-of-the-tongue (TOT) state. Yet, older adults report fewer competing alternatives despite experiencing more TOTs (Burke et al., 1991). So, while the domain general nature of the Inhibition Deficit Hypothesis accounts for age-related declines in memory (May, Zacks, Hasher, & Multhaup, 1999), postural control (Redfern, Jennings, Martin, & Furman, 2001), and some phonological neighborhood density effects (Chen & Mirman, 2012; Luce & Pisoni, 1998; Vitevitch & Luce, 2016), there are domain specific aspects of cognitive aging which remain unaccounted for by the Inhibition Deficit Hypothesis (Burke, 1997).

Neural Aging Theories

Beyond behavioral studies, investigations utilizing functional magnetic resonance imaging (fMRI) have also been used to probe mechanisms underlying cognitive aging. Functional MRI studies have found mixed patterns of activation differences in older adults. Briefly, older adults engage task-relevant regions less (Cabeza, 2002) and, concurrently, engage more anterior(Davis, Dennis, Daselaar, Fleck, & Cabeza, 2008) and bilateral activation (Cabeza, 2002) compared to younger participants. This has led to the development of two broad classes of neural aging theories—dedifferentiation and compensation theories. Dedifferentiation theories suggest that with age, neural processing naturally becomes more diffuse—and consequently less effectual—which leads to poorer performance (Li, Lindenberger, & Sikström, 2001). Compensation theories of neural aging say that processing becomes less efficient with age and, as a result, older adults engage more diffuse regions to maintain behavioral performance

(Cabeza, 2002). The Compensation and Related Utilization of Neural Circuits Hypothesis or CRUNCH provides a third, hybrid class of neural aging theories; CRUNCH suggests individuals engage in additional activation to maintain behavioral performance but, at a certain level of deficit, behavioral compensation cannot be achieved. Thus, if CRUNCH holds, older adults should show additional activation on tasks because tasks are more difficult for older adults than younger adults (Reuter-Lorenz & Cappell, 2008).

Importantly, dedifferentiation, compensation, and CRUNCH are domain general theories of neural aging which do not specifically predict activation to a language production task. However, interpretation of functional results in the context of these theories can inform theoretical accounts of age-related language production differences. For instance, findings from language studies indicate older adults engage cognitive control mechanisms as a compensatory mechanism when faced with challenging situations (Shafto & Tyler, 2014). As such, these findings point to the role of domain general resources—like inhibition—in driving language production ability, and support the Inhibition Deficit Hypothesis. Rizio and colleagues, on the other hand, found a CRUNCH pattern in their study of older adults' naming ability; they found that older adults engaged in less activation to a phonological task but more activation in a semantic task. However, in their study, older adults showed equivalent reaction time but poorer accuracy compared to younger adults (Rizio et al., 2017), suggesting that behavioral compensation may not always be possible.

White Matter

While it is clear that older and younger adults have differences in behavioral outcomes and functional activation patterns, the underlying biological substrates driving these differences remain in question. One potential aspect of aging which may drive cognitive and functional changes is white matter integrity decline. White matter integrity is measured with a Magnetic Resonance Imaging technique called Diffusion Tensor imaging (DTI). DTI measures the amount and apparent direction of water diffusion in the brain (Jones, Knösche, & Turner, 2013). White matter structures, as well as other factors, force the otherwise Brownian diffusion of water molecules into a more anisotropic pattern of diffusion (Jones et al., 2013). Thus, modeling the shape of diffusion at each voxel of the brain can provide an indirect measure of the integrity of white matter fibers at that point.

White matter integrity reliably decreases with age (Bennett & Madden, 2014; Head et al., 2004) and has been found to mediate brain-behavior relationships in a host of cognitive tasks. For instance, Fjell and colleagues found that 82.5% of variance in longitudinal declines in executive function could be accounted for by white matter volume loss (2017). Hedden et al. found that white matter integrity was a significant predictor of processing speed and executive function (2016). Hedden et al. also found that age significantly mediated these effects, pointing to white matter as a significant factor in age-related cognitive differences (2016). Further, Gaze and colleagues found that better white matter integrity along specific subsets of white matter tracts significantly predicted faster processing speed, better reasoning ability, and better episodic memory but found no subset of tracts related to vocabulary ability (2016).

Notably, white matter integrity has been shown to decrease along an anterior-to-posterior gradient as is expected with the Inhibition Deficit Hypothesis (Davis et al., 2009; Pfefferbaum, Adalsteinsson, & Sullivan, 2005). White matter has also been found to decline diffusely across the language network as would be expected in the Transmission Deficit Hypothesis (Stamatakis et al., 2011). Stamatakis and colleagues found that white matter along the Superior Longitudinal Fasciculus (SLF), Inferior Longitudinal Fasciculus (ILF), genu, anterior corpus callosum, posterior internal capsule, middle cingulum, corticospinal tract, and the corticobulbar tract are positively correlated with ability to provide proper names of famous individuals (Stamatakis et al., 2011). This suggests that white matter integrity along language-network tracts is an important factor in naming ability. Moreover, Stamatakis and colleagues found that word finding failures were only correlated with white matter integrity along the posterior SLF (2011), which traverses regions associated with phonological processing (Hickok & Poeppel, 2007). Therein, the work of Stamatakis and colleagues provides evidence toward the Transmission Deficit Hypothesis. Conversely, using principle components analysis, Rizio and Diaz found that white matter integrity along the Frontal Aslant Tract (FAT) and SLF/ arcuate (AF) predicts working memory (Rizio & Diaz, 2016; Rizio et al., 2017). This empirically supports the hypotheses set forth by Dick, Bernal, and Tremblay that the SLF/AF and FAT are both relevant to language production ability; the SLF subserves phonological processing while the FAT subserves the use of executive resources during speech production (2014). Thus, there is preliminary evidence to suspect that both the Transmission Deficit Hypothesis and the Inhibition Deficit Hypothesis could be tested through examination of white matter integrity in older adults.

Two measures of white matter integrity are of particular relevance for the study of age-related changes in white matter; Fractional Anisotropy (FA) and Radial Diffusivity (RD). FA represents a summary measure of how directional or anisotropic water diffusion patterns are. Given that more intact white matter should force water to be more anisotropic, higher directionality and higher FA indicate better white matter integrity. RD is the converse. RD represents how much water is diffusing in the direction radial or perpendicular to white matter tracts. If more water is diffusing in this direction, higher RD scores indicate poorer tract integrity. Though both FA and RD were used in this study, RD has been shown to be more sensitive to age-related changes in white matter integrity when compared to FA (Davis et al., 2009).

Chapter 2

The Present Study

The present study aims to clarify the role of white matter in contributing to either age-related differences or similarities in phonological and semantic language systems, respectively. To accomplish this, I tested the relationship between white matter integrity in younger and older adults and behavioral outcomes (i.e., accuracy) on a task of language production called the Picture Word Interference (PWI) task as well as on the Stroop task of inhibition. In the PWI task, participants named a picture while ignoring either phonological or semantic distractors. In addition to examining behavioral outcomes, this study examined the relationship between white matter and functional activation to the PWI task in regions of interest related to the phonological and semantic conditions (Rizio et al., 2017). In the Stroop task, participants named the print color of the word presented while inhibiting the reading of the written word.

Given the predictions of the Dual Stream Model of language (Hickok & Poeppel, 2007) and the evidence that the SLF/AF and FAT are related to language processing (Dick et al., 2014; Rizio et al., 2017) while the ILF and MDLF are related to language comprehension (Dick et al., 2014), this study conceptualized the SLF/AF and FAT as phonological tracts and the ILF and MDLF as semantic tracts. The Fronto-striatal tract (FS) has been shown to be related to inhibition (Liston et al., 2006). As such the FS was conceptualized as a cognitive control tract. In order to ascertain the role of white matter along the dorsal (phonological), ventral (semantic), and FS (cognitive control) tracts, analyses tested specific hypotheses which align with the predictions of either Transmission Deficit (1) or Inhibition Deficit Hypotheses (2).

Specifically, this analysis compared the two measures of white matter integrity from along these tracts, FA and RD, to accuracy on the PWI task and Stroop Effect scores on the

Stroop task. No reaction time measures were used because, in initial analyses done by Rizio and colleagues, there were no significant differences in reaction time between older and younger adults (Rizio et al., 2017). I also compared white matter measures to activation in regions informed by the findings of Rizio et al., who showed that, in the phonological condition, younger adults had significantly stronger activations in the right post central gyrus, right supramarginal gyrus, and bilateral middle temporal gyrus while older adults showed no regions of stronger activation than younger adults; during the semantic condition, younger adults showed no regions of stronger activation than older adults, but older adults showed stronger activations than younger adults in the left frontal gyrus, middle posterior cingulate gyrus, left superior parietal lobe, bilateral and middle precuneus, bilateral lingual gyrus, left occipital fusiform gyrus, right cuneus, and left cerebellum (Rizio et al., 2017). This study performed analyses with these significant regions as pre-defined ROIs to further characterize the interplay between white matter integrity and functional activation to the PWI task in older and younger adults.

Transmission Deficit Hypothesis Predictions (1)

1.1 Structure-Behavior Predictions

The Transmission Deficit Hypothesis predicts that better behavioral accuracy will be related to better white matter integrity along dorsal (SLF/AF and FAT) but not ventral (ILF and MLDF) tracts. Though the Transmission Deficit Hypothesis specifies this for older adults, to preserve statistical power, this study first investigated main effects (across participants) before looking for age interactions. As such I predicted that across participants:

> 1.1.A Better dorsal tract integrity would be associated with better accuracy on the phonological condition of the PWI task but not the semantic condition.

1.1.B Ventral tract integrity would not be associated with accuracy on the semantic or phonological conditions on the PWI task.

I further predicted that the effect size of the relationship between dorsal tract (SLF/AF and FAT) integrity and phonological accuracy (1.1.A) would but would be statistically larger than any (significant or insignificant) relationship between ventral tract integrity (ILF and MDLF) and semantic accuracy (1.1.B). In other words I predicted:

1.1.C The relationship between white matter integrity along dorsal tracts and phonological accuracy would be stronger than the relationship between ventral tract integrity and semantic accuracy.

1.2 Structure-Activation Predictions

Similar to the behavioral predictions, the Transmission Deficit Hypothesis would predict lower white matter integrity would manifest in significant activation outcomes in phonological but not semantic conditions. Again, this prediction is made specifically for older adults but, to preserve statistical power, main effects were tested before looking for age-specific effects. Because the previously published results of this study indicated greater activation in the phonological condition (Rizio et al., 2017) and better performance in younger adults (Rizio et al., 2017), I expected better white matter integrity to be related to better performance. I hypothesized:

1.2.A Better dorsal tract integrity would be associated with stronger activation in the phonological ROI but not the semantic ROI.

Again, since the Transmission Deficit Hypothesis predicts no relationship between white matter integrity along ventral tracts and outcomes of a semantic task, I expect:

1.2.B Semantic tract integrity would not be related to activation in the phonological or semantic ROI.

Analogous to behavioral predictions, I expect the effect size of the relationship between dorsal tract integrity and phonological activation (1.2.A) would be statistically larger than any (significant or insignificant) relationship between ventral tract integrity and semantic activation (1.2.B). In other words:

1.2.C The relationship between white matter integrity along dorsal tracts and phonological activation would be stronger than the relationship between ventral tracts and semantic activation.

1.3 Interaction and Mediation Predictions

Given that the predictions of the Transmission Deficit Hypothesis are made specifically to account for age-related differences in language production ability, I also hypothesize age by white matter interaction effects. That is, I expected that the previously hypothesized white matter-accuracy and white matter-activation relationships would appear stronger in older adults because younger adult white matter and accuracy should be at ceiling. As such, I hypothesized:

1.3.A There would be an age x dorsal white matter integrity interaction such that the relationship between accuracy on the phonological condition and white matter along dorsal tracts is stronger in older compared to younger adults.

1.3.B There would be an age x dorsal white matter integrity interaction such that the relationship between activation to the phonological condition and white matter along dorsal tracts is stronger in older compared to younger adults.

Moreover, if white matter is indeed the neural mechanism underlying age-related deficits in language production ability, I expect that white matter integrity would statistically mediate age-related differences in behavioral and functional outcomes such that:

1.3.C White matter integrity along the dorsal tract would mediate the relationship between age and performance on the phonological condition of the PWI task

1.3.D White matter integrity along the dorsal tract would mediate the relationship between age and activation to the phonological condition of the PWI task

Inhibition Deficit Hypotheses Predictions (2)

2.1 Structure-Behavior Predictions

The Inhibition Deficit Hypothesis suggests that declines in inhibition account for declines in other abilities such as language (Hasher & Zacks, 1988). As such, I expected better white matter integrity along a tract known to be related to cognitive control, the FS, would predict better performance on a behavioral task of inhibition (the Stroop effect measure of the Stroop task) as well as on all conditions of the PWI task. Though these predictions are, again, specifically made to account for cognitive differences between younger and older adults, main effects are examined first. Thus, I predicted across participants:

2.1.A Better white matter integrity along the FS would predict lower Stroop effect scores.

2.1.B Better white matter integrity along the FS would predict better accuracy on the phonological condition of the PWI task.

2.1.C Better white matter integrity along the FS would predict better accuracy on the semantic condition of the PWI task.

2.2 Structure-Activation Predictions

Since the Inhibition Deficit Hypothesis suggests that greater inhibition ability would be associated with better performance and because, in the sample, the younger adults who drive activation patterns in the phonological ROI were also better performers (Rizio et al., 2017), I expect that across participants:

2.2.A Better FS tract integrity would be associated with stronger activation in the phonological ROI.

Conversely, in this sample, older adults showed stronger activation in the semantic ROI, were worse performers on the semantic condition (Rizio et al., 2017), and were expected to have lower white matter integrity. Thus, I expect that across participants:

2.2.B Better FS tract integrity would be associated with weaker activation in the semantic ROI.

2.3 Interaction and Mediation Predictions

Given that the predictions of the Inhibition Deficit Hypothesis are made specifically to account for age-related differences in language production ability, I also hypothesize age by white matter interaction effects. That is, I expected the previously hypothesized FS integrity-accuracy and FS integrity–activation relationships would appear stronger in older adults because younger adult white matter integrity and performance should be at ceiling. As such I expected:

2.3.A There would be an age x white matter integrity interaction in the Stroop task such that the relationship between Stroop Effect and FS tract integrity is stronger in older relative to younger adults.

2.3.B There would be an age x white matter integrity interaction in the phonological condition of the PWI task such that the relationship between accuracy and FS tract integrity is stronger in older relative to younger adults.

2.3.C There would be an age x white matter integrity interaction in the semantic condition of the PWI task such that the relationship between accuracy and FS tract integrity is stronger in older relative to younger adults .

In addition, if white matter is the neural mechanism underlying age-related deficits in language production ability, I would expect that FS white matter integrity would statistically mediate age-related differences in behavioral outcomes such that:

2.3.D White matter integrity along the FS would mediate the relationship between age and Stroop effect scores.

2.3.E White matter integrity along the FS would mediate the relationship between age and accuracy on the phonological condition of the PWI task.

2.3.F White matter integrity along the FS would mediate the relationship between age and accuracy in the semantic condition of the PWI task.

Chapter 3

Methods

The Picture Word Interference Study

Data for this study came from the Picture Word Interference study conducted by the Language and Aging Lab at the Pennsylvania State University. At the time this study began, all behavioral, structural, and functional data had been collected and cleaned; Demographic statistics on the participants were known; basic descriptive statistics such as group means, standard deviations, and between-group-differences for the behavioral and functional data were known; and tracking of the FAT had also been completed. However, no analysis had yet been performed on the white matter integrity data. All other tracts (SLF, ILF, MDLF, and FS) were modeled over the course of the experiment using procedures modified from those used by Rizio and Diaz (2016). Before this study began, exclusion masks were already defined for the SLF, ILF, and MDLF. These were unmodified. Seeds, targets, and waypoints, for the SLF, ILF, and MDLF were theoretically defined in the procedure used by Rizio and Diaz (2016) but these aspects were modified throughout the tracking process (see Appendix A for final procedures). Though Tractography procedures were modified throughout the study to optimize tracking, no DTI data was extracted or analyzed until after the tractography procedure for that tract was finalized to minimize researcher bias.

Participants

20 younger (18-31, females=10, mean age=23.7) and 20 older (60-79, females=15, mean age= 67) adults participated in this experiment. All were healthy, right-handed, monolingual English speakers with normal or corrected-to-normal vision as measured by the Freiburg Visual Acuity and Contrast Test (Bach, 1996). No participants reported a history of neurological or

psychological disorders. All participants scored at least 27 on the Mini Mental State Examination (MMSE; (Folstein, Folstein, & McHugh, 1975). All participants provided written informed consent, and all experimental procedures were approved by the Institutional Review Board at the Pennsylvania State University.

Procedure

Before the MRI session, participants completed assessments to determine handedness and language history. Participants also completed a neuropsychological battery to assess speed, inhibition, working memory and language ability. Prior to scanning participants practiced overt picture naming in a simulation scanner both to familiarize participants with the MR environment as well as to allow for practice minimizing head movement. Importantly, participants were not familiarized with target words or picture before entering the MRI. During the MRI session, overt verbal responses were recorded and filtered using an MR-compatible fiber optic microphone system (Optoacustics Ltd.,Or-Yehuda, Israel).

Stimulus Materials

Stimuli for the picture word interference paradigm consisted of 240 colored images of common, concrete nouns such as animals, clothing, food, and household items from two normed picture databases (Brodeur, Guérard, & Bouras, 2014; Moreno-Martínez & Montoro, 2012) with a superimposed written distractor word (See Figure 1). Average name-agreement of

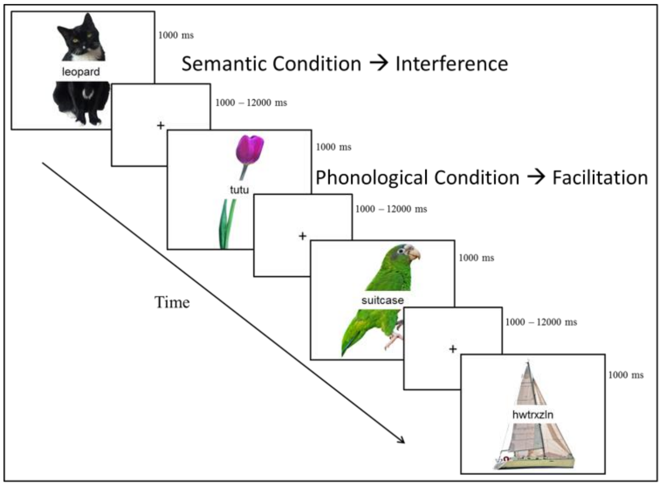

Figure 1 In the Picture Word Interference (PWI) Task, participants are instructed to name the picture and ignore the superimposed word which was either semantically or phonologically related to the word or an unrelated word or letter string. This study will examine only the semantic and phonological conditions. Figure partially reproduced with permission from (Rizio et al., 2017).

pictures, as taken from the aforementioned databases, was 72%. The average log-transformed HAL frequency of words 7.48 (SD=1.91) on a scale of 0-17. The range of frequencies was intentionally large as another aim of this work, reported elsewhere, was to assess the influence of frequency on picture naming (Rizio et al., 2017).

Name-agreement of pictures, lexical frequency of distractors, and length of the distractors were matched across four distractor conditions; semantic (categorical), phonological, unrelated, and, non-word. Each condition contained 60 unique pictures and distractor words with no repetition within subjects. However, for each of the 240 target images, four different distractors were created, resulting in 4 stimuli lists. Across participants, each target was paired with each condition, ensuring that item effects did not drive any statistical differences between conditions (Rizio et al., 2017).

Table 1. Stimuli Word Characteristics

	Log Frequency Mean (SD)	Word Length Mean (SD)	Target-Distractor categorical relatedness Mean (SD)	Target-Distractor orthographic similarity Mean (SD)
Target	7.48(1.91)	6.47(2.07)	--	--
Semantic Condition	7.43(1.66)	6.78(1.92)	6.16(1.37)	0.10(0.09)
Phonological Condition	7.70(1.86)	6.13(1.80)	1.36(0.40)	0.49(0.11)

Across semantic and phonological conditions, lexical and orthographic characteristics were matched. Data partially reproduced with permission from (Rizio et al. 2017).

In the semantic condition distractors were semantically, but not associatively, related to the targets. Semantic relatedness was measured on a scale of one (not at all categorically related) to seven (very categorically related) by 47 younger adults who did not participate in the MRI session. There was a significant difference in judgment of categorical relatedness across conditions. Specifically, semantic distractors were statistically more related to their respective targets than were the phonological and unrelated distractors (Rizio et al., 2017)

Phonologically related distractors shared at least two initial phonemes with their respective target and were matched across lists on initial sound and stress as defined by the Carnegie Mellon Pronouncing Dictionary. Phonological relatedness was quantified by calculating the orthographic similarity (OS) of target and distractor pairs. OS was calculated in NIM (Guasch, Boada, Ferré, & Sánchez-Casas, 2013) using the following equation: Orthographic Similarity = Graphemic Similarity between the target and distractor/Graphemic Similarity of the target with itself. OS values can range from 0 (no similarity) to 1 (identical). There was a significant difference among the OS values across conditions. Specifically, OS values were higher in in the phonological condition than in the semantic, unrelated, and non-word conditions (Rizio et al., 2017).

Unrelated distractors were neither categorically nor phonologically related to the target. Non-word distractors consisted of random consonant strings that did not start with the same letters as the target. Each trial consisted of a target picture presented on a white background (7 × 5.5 inches) and a distractor word superimposed upon the center of the picture. Pictures were constrained to either 7’’ wide or 5.5’’ tall to regularize the size of the objects without distorting the aspect ratio and distractors were presented in Courier New 18 point font. Target pictures and distractor words were presented simultaneously (i.e., SOA = 0) for 1 s. Participants were instructed to name the target picture, but ignore the distractor word, and to respond as quickly as possible while still responding accurately. A fixation cross was presented between each stimulus presentation [interstimulus interval (ISI) range = 1–12 s, average ISI = 4 s]. ISIs were optimized with Optseq2 (Dale, 1999). Participants were instructed that they had both the duration of the target presentation, as well as the duration of the ISI, to make their response. Trials were randomized so that no more than three of the same distractor condition appeared in a row. Each of four runs (315 s) began and ended with the presentation of a fixation cross.

Neuropsychological Measures

Before entering the scanner participants completed a neuropsychological battery which included tests of verbal fluency, non-verbal working memory, forward and backward digit span, verbal memory, and vocabulary knowledge. Simple speed of processing was assessed using a button press task and complex speed of processing was assessed using the Digit Symbol Task. Data on participants reading habits were gathered using the Author and Magazine Recognition Tests and the Comparative Reading Habits Questionnaire. Finally, participants completed the Stroop task of Executive Function which was utilized in this study.

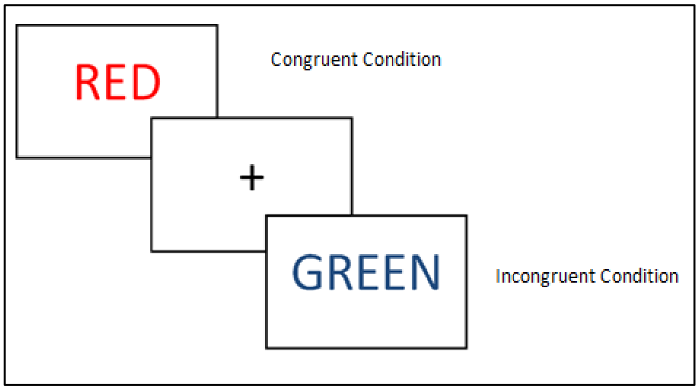

Figure 2. In the Stroop Task of Executive Function participants are asked to report the text color of the word presented rather than reading the word. In congruent conditions text colors match the word presented; in incongruent conditions the color and word do not match.

For the purposes of this study, analyses focused on the Stroop Effect measure of the Stroop task. That is, in the Stroop task, participants are asked to report the color in which a word was printed. In congruent trials, the color of the print is the same as the word which is printed (e.g., red printed in red) while in the incongruent condition, the word printed is the name of a color which is not the same as the color in which the word is printed (e.g., green printed in blue; see Figure 2). In order to report the color in incongruent conditions, participants must engage inhibition to a greater degree than in the congruent condition. Any increases in reaction time, known as the Stroop Effect, provide a measure of participants' inhibition ability. Stroop Effect is calculated by subtracting the reaction time on congruent trials from the reaction time on incongruent trials. Thus, lower Stroop Effect scores are indicative of better performance.

Acquisition of MRI data

Data were acquired using a 3 T Siemens Prisma Fit MRI scanner with a 20-channel head coil. First, 3 sequential sagittal 0.5x0.5x7mm slices were collected in the anterior to posterior direction with a FOV= 250 mm; TR=8.6ms TE=4.0 ms; 7ms echo spacing, and flip angle = 20^o to ensure the participant was properly positioned in the FOV. Following this localizer, anatomical images and resting data were acquired, followed by four functional runs, collection diffusion data, and field map collection.

Functional Data

An advanced shim was applied before the first of four functional runs. Functional images were collected using an echo-planar imaging (EPI) sequence along the axial plane with an anterior to posterior phase encoding direction and 0.49 ms echo spacing. Forty-one interleaved slices were acquired with a voxel size of 3 mm^3, FOV = 240 mm^2; matrix = 80 mm^2. The TR = 2500 ms; TE = 25.0 ms, and the flip angle = 90˚. Fat saturation was used. A total of 128 volumes per functional run were collected however two volumes were deleted at the beginning of each functional run, leaving 126 volumes per functional run for analysis. 41 interleaved contiguous slices of resting state data were also collected using the same parameters.

Brain matter was segmented from non-brain matter using optiBET (Lutkenhoff et al., 2014), and quality assured the data for motion and artifacts using software. Within-run and within-person analyses of activation were conducted using FSL FEAT (Smith et al., 2004). To assess the relationship between white matter integrity and activation, mean activation of the top

10% of voxels was extracted from two regions of interest (ROIs)—a phonological and semantic region. The phonological ROI was defined as the voxels significantly more activated by younger adults during the phonological task. In total, the phonological ROI was comprised of 2486 voxels in a region of the left middle temporal gyrus at the coordinates x=-64, y=-48 and z=0 (See figure 3).

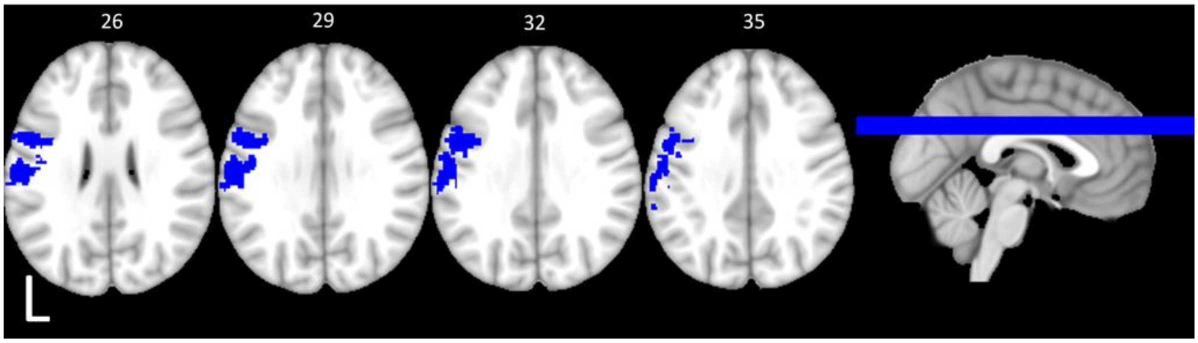

Figure 3. The phonological ROI used for the structure-activation analyses. These regions were those significantly more activated by younger adults compared to older adults during the phonological conditon of the PWI task and encompass portions of the left middle temple gyrus.

The semantic ROI was defined as the voxels significantly more activated by older adults during the semantic task which resulted in a region comprised of 382 voxels of the left frontal pole at the coordinates x=-28, y=38, z=46 and a separate 16 voxels in the left middle frontal gyrus at the coordinates x=-34,y-6, z=40 (See figure 4).

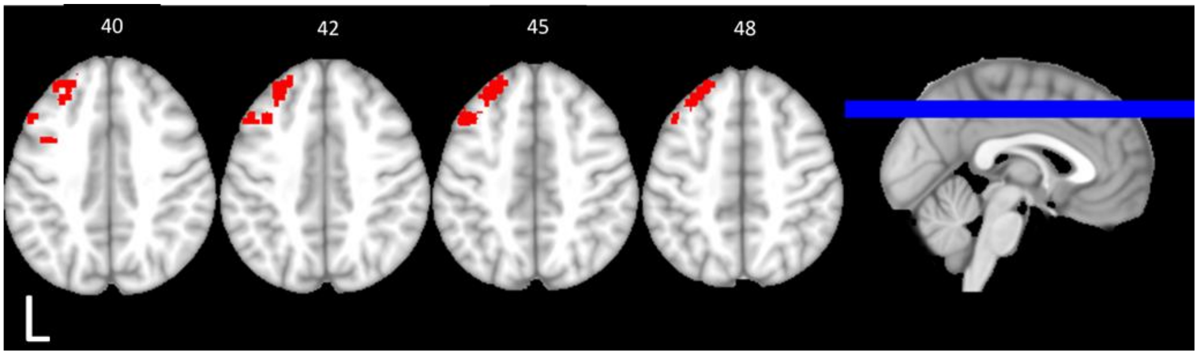

Figure 4. The semantic ROI used for the structure-activation analyses. These regions were those significantly more activatied by older adults compared to younger adults during the semantic conditon of the PWI task and encompass portions of the left middle frontal gyrus.

Anatomical Images

T-1 weighted anatomical images were collected using a single shot magnetization-prepared rapid acquisition gradient echo (MP RAGE) sequence along the sagittal plane with an anterior to posterior phase encoding direction, and 7 ms echo spacing. A PE acceleration factor of R= 2 was applied using a GRAPPA technique with 24 reference lines.160 contiguous slices, 1mm thick, were acquired in ascending order with a voxel size of $1mm^3$, FOV = 256 mm^2, and matrix = 256 mm^2. The TR = 2300 ms; TE = 2.28 ms; TI = 900 ms; and the flip angle = 8˚. Fat suppression was not used.

Diffusion weighted data were acquired from 36 directions along the axial plane with a posterior to anterior phase encoding direction. A PE acceleration factor of R= 2 was applied using a GRAPPA technique with 40 reference lines. 62 interleaved contiguous slices were acquired with a voxel size 2 mm^3, FOV = 240 mm^2; matrix = 80 mm^2. The TR = 1000 ms; TE = 89 ms, .78 ms echo spacing, and the flip angle = 90˚. Fat saturation was used.

Diffusion weighted data were quality assured for motion and artifacts using DTIprep. The diffusion tensor model was fit to the data using DTIfit and crossing fibers were modeled using BedpostX. Six separate tracts were then modeled probabilistically using ProbtrackX2 in FSL. The tracts included the Superior Longitudinal Fasiculus-III (SLF)/ acruate (AF), the Frontal Aslant Tract (FAT), the Middle Longitudinal Fasiculus (MDLF), the Inferior Longitudinal Fasiculus (ILF), and the Frontostriatal Tract (FS; see figure 5). Each tract was model separately in each hemisphere, in each participant, but only left hemisphere data was used for this analysis. These models were then used as masks to extract two measures of white matter integrity; mean Fractional Anisotropy (FA) and mean Radial Diffusivity (RD) along each tract.

Because there are theoretical relationships, as well as statistical collinearity among tracts, composite scores were created to reduce the number of comparisons being made. Specifically, SLF/AF and FAT tract values were averaged together to create two composite measure of dorsal tract integrity—dorsal FA and dorsal RD—for each hemisphere in each participant. Likewise MDLF and ILF tract integrity values were averaged together to create the composite measures, ventral tract FA and ventral tract RD for each hemisphere, in each participant.

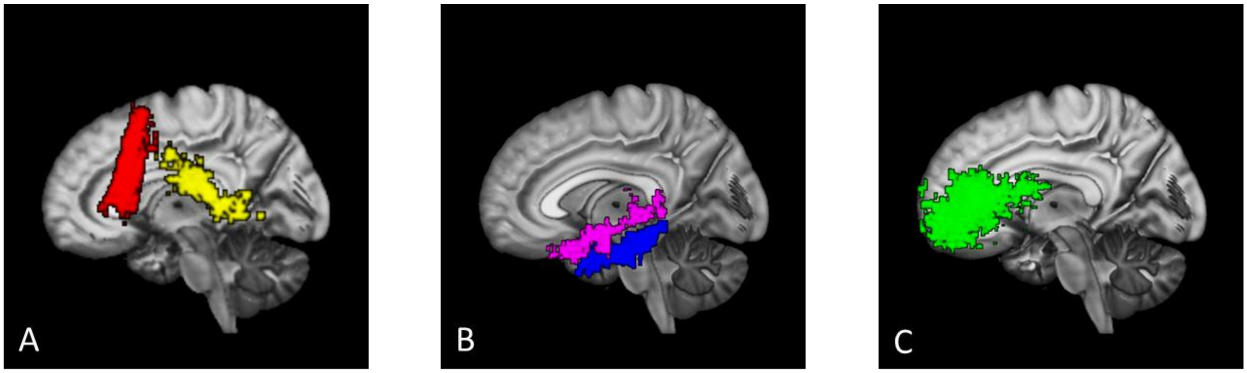

Figure 5. The above are representative examples of the white matter tracts modeled at the individual level in this study. Box A displays an example of the each of the dorsal tracts; the FAT (Red) and the SLF (Yellow). Box B displays an example of each of the ventral tracts; the MDLF (Purple) and the ILF (Blue). Box C. displays an example of the FS (Green). All tracts were identified reliably in 39 out of 40 participants.

Data Analysis

Structure-Behavior Relationships

To assess the association between tract integrity and performance on the PWI task, bivariate correlations were used to test the relationships between white matter integrity along dorsal tracts, ventral tracts, and the FS tract and performance on phonological and semantic conditions of the PWI task. One-tailed Fisher R-Z transformations were then used to assess if any of the tracts (dorsal, ventral, or FS) were significantly more related to behavioral performance on the PWI than the other tracts.

To assess the association between tract integrity and inhibition, bivariate correlations tested the relationships between white matter integrity along the dorsal, ventral, and FS tracts and performance on the Stroop Effect condition of the Stroop task. Again, one-tailed Fisher R-Z transformations assessed if any of the tracts were significantly more related to performance on the Stroop task than the other tracts

Structure-Activation Relationships

Bivariate correlations between activation in semantic and phonological ROIs and the dorsal, ventral, and FS tracts tested for significant structure-activation relationships. One-tailed Fisher R-Z transformations assessed if any of the tracts were significantly more related to level of activation the other tracts

Interaction and Mediation Analyses

Multiple regression models predicting performance on the PWI task from age, white matter integrity, and the interaction of these factors were built to test for differences in the effect of white matter integrity across age. However in order to minimize statistical tests, the interaction was only tested if multiple regression models with the two principle factors (age and white matter integrity) indicated that both predictors were significant. Importantly, separate regression models were built for each dorsal, ventral, and FS tracts. The same model building procedure was used to test for an interaction between white matter and age in predicting Stroop Effect and mean activation in the phonological and semantic ROIS.

Mediation analyses were conducted using conceptual model 4 in the PROCSS macro in SPSS(Hayes, 2016). White matter was hypothesized to mediate two types of relationships: age-behavior and age-activation. To minimize comparisons, mediation analyses were only conducted if two conditions were met; first, a significant age-behavior or age-activation relationship was present to be mediated; second, a significant bivariate relationship existed between age and white matter integrity along the relevant tract (i.e., the tract hypothesized as the mediator).

Chapter 4

Results

A one-way between groups ANOVA revealed that younger adults were more accurate compared to older adults on both phonological (*F*(1,36) = 7.21, *p* = .01) and semantic (*F*(1,36) = 7.07, *p* = .01) trials of the PWI task but there were no significant differences in reaction time on either phonological (*F*(1,36)= .33, *p*= .57) or semantic (*F*(1,36) = .13, *p* = .72) trials. Stroop Effect scores were also significantly worse in older adults (*F*(1,36) = 24.32, *p* < .001). In the phonological condition of the PWI task, younger adults had stronger activation than older adults in the phonological ROI (*F*(1,36) = 19.21, *p* < .001). In the semantic condition, younger adults had weaker activation in the semantic ROI than older adults (*F*(1,36) = 13.09, *p* = .001). Older adults also had poorer white matter integrity than younger adults along all tracts (See table 2).

Table 2. Average Participant Behavioral Performance and White Matter Integrity

YA=18, OA=20	Mean		SD	
Behavioral	YA	OA	YA	OA
Phonological Accuracy (% correct)	**.77***	**.68***	.07	.11
Semantic Accuracy (% correct)	**.72***	**.64***	.09	.10
Stroop Effect (ms)	**9.13****	**92.87****	30.50	65.91
Functional Activation				
Phonological	**.18****	**.05****	.09	.09
Semantic	**.08****	**.23****	.11	.14
White Matter Intgerity				
Dorsal FA	**.45****	**.42****	.01	.02
Ventral FA	**.44****	**.42****	.01	.02
FS FA	**.43****	**.40****	.01	.02
Dorsal RD	**5.50x10^{-4}****	**5.97x10^{-4}****	.16x10^{-4}	.38x10^{-4}
Ventral RD	**5.77x10^{-4}****	**6.00 x10^{-4}****	.17x10^{-4}	.30x10^{-4}
FS RD	**5.51 x10^{-4}****	5.91 **x10^{-4}****	.14x10^{-4}	.34x10^{-4}

**Difference between YA and OA at p<.05*

***Difference between YA and OA significant at p<.005*

Transmission Deficit Hypothesis Results

Structure-Behavior Relationships

1.1.A Better white matter integrity along dorsal tracts was associated with greater accuracy on the phonological condition of the PWI task as measured by RD (r(36) = -.35, *p* = .03) and FA (r(36) = .41, *p*= .01). Accuracy on Semantic trials was not significantly related to dorsal tract RD (r(36) = -.18, *p* = .29) or FA (r(36) = .17, *p* = .30).

1.1.B Ventral white matter integrity as measured by RD was not significantly related to phonological accuracy (r(36) = -.28, *p* = .10) or semantic accuracy (r(36) = .12, *p* = .48). Ventral FA was also not significantly related to phonological accuracy (r(36) = -.23, *p* = .17) or semantic accuracy (r(36) = .01, *p* = .98).

1.1.C One-tailed Fisher R-Z transformation tests reveal that the correlation between phonological accuracy and dorsal RD was statistically stronger than the relationship between semantic accuracy and ventral RD (z = -2.05, *p* = .02). Similarly, the relationship between phonological accuracy and dorsal FA was statistically stronger than the relationship between semantic accuracy and ventral FA (z = 1.85, *p* = .03).

Structure-Activation Relationships

1.2 Better dorsal white matter integrity predicted stronger activation in the phonological ROI when white matter integrity was measured by RD (r(36) = -.42, *p*=.008) and FA (r(36) = .46, *p* = .004). Dorsal white matter integrity was not significantly related to activation in the semantic ROI when white matter integrity was measured by RD (r(36) =

.30, *p* = .06). However dorsal FA did predict weaker activation in the semantic ROI (r(36) = -.35, *p* = .03).

1.2.B Ventral RD was not significantly related to activation in the semantic ROI (r(36) = .16, *p* = .34) or phonological ROI (r(36) = -.02, *p* = .89) and ventral FA was not significantly related to activation in the semantic ROI (r(36) = -.21, *p* = .21) or activation in the phonological ROI (r(36) = .19, *p* = .27).

1.2.C The strength of relationship between dorsal RD and phonological activation was statistically the same as the strength of relationship between ventral RD and semantic activation (z = 1.21, *p* = .11). Similarly, the strength of the relationship between the dorsal FA and phonological activation was the same as the relationship between ventral FA and semantic activation (z= 1.2, *p* = .12).

Interaction and Mediation Analyses

1.3.A In a multiple regression model including only dorsal RD and participant age to predict phonological accuracy, neither dorsal RD (β = -.15, $r_{rdaccuracy.age}$ = -.13, *t*(35) = -.76, *p* = .45) nor age (β = -.32, $r_{rdage.accuracy}$ = -.27, *t*(35) = -1.65, *p* = .11) was a significant predictor of phonological accuracy. Therefore, no interaction was tested. Further, there was no test of the interaction between dorsal FA and age as, in a model containing these variables with no interaction term, neither age (β = -.25, $r_{faaccuracy.age}$ = .16, *t*(35) = -1.12, *p* = .27) nor dorsal FA (β = .22, $r_{faage.accuracy}$ = -.19, *t*(35) = .97, *p* = .34) were significant predictors of phonological accuracy.

1.3.B In a multiple regression model including only dorsal RD and participant age to predict phonological activation, age was a significant predictor ($\beta = -.54$, $r_{rdage.activation} = -.46$, $t(35) = -3.04$, $p = .004$) but dorsal RD was not ($\beta = -.08$, $r_{rdage.activation} = -.08$, $t(35) = -.47$, $p = .64$). Therefore, the interaction of dorsal RD and age was not tested. In a model containing only dorsal FA and age as predictors of phonological activation, age remained a significant predictor ($\beta = -.55$, $r_{faage.activation} = -.41$, $t(35) = -2.67$, $p = .01$) but dorsal FA was not ($\beta = .05$, $r_{faactivation.age} = .04$, $t(35) = .25$, $p = .81$). Therefore, the interaction of dorsal FA and age was not tested.

1.3.C A mediation analysis using a bootstrap approach with 5000 samples revealed that after controlling for the effect of dorsal RD, the direct effect of age on phonological accuracy was no longer significant ($\beta = -1.5\text{x}10^{-3}$, $t(36) = -1.65$, $p = .11$) despite an insignificant indirect effect of dorsal RD ($\beta = .40\text{x}10^{-3}$, SE = $.60\text{x}10^{-3}$, 95% CI = $-1.8\text{x}10^{-3}$, $.6\text{x}10^{-3}$). This suggests that RD is not a t mediator but is a moderator. Similarly, after controlling for the effect of dorsal FA, the direct effect of age on phonological accuracy was no longer significant ($\beta = -1.1\text{x}10^{-3}$, $t(36) = -1.12$, $p = .27$) despite an insignificant indirect effect of dorsal FA ($\beta = -.70\text{x}10^{-3}$, SE = $.80\text{x}10^{-3}$, 95% CI = $-2.40\text{x}10^{-3}$, $.70\text{x}10^{-3}$). This suggests that dorsal FA is not a mediator but is a moderator of the relationship between age and phonological accuracy.

1.3.D A mediation analysis using a bootstrap approach with 5000 samples revealed that after controlling for the effect of dorsal RD, the direct effect of age on phonological activation remained significant ($\beta = -2.7\text{x}10^{-3}$, $t(36) = -3.04$, $p = .005$) and the indirect effect of dorsal RD was insignificant ($\beta < -.3\text{x}10^{-3}$, SE = $.07\text{x}10^{-3}$, 95% CI = $-1.7\text{x}10^{-3}$, $1.0\text{x}10^{3}$). This suggests that dorsal RD is not a mediator or moderator of the relationship between age and phonological activation. After controlling for the effect of dorsal FA, the direct effect of age on phonological activation remained significant ($\beta = -2.70\text{x}10^{-3}$, $t(36) = -2.69$, $p = .01$) despite an insignificant indirect effect of dorsal FA ($\beta = -0.20\text{x}10^{-3}$, SE = $0.80\text{x}10^{-3}$, 95% CI = $-1.80\text{x}10^{-3}$, $1.40\text{x}10^{-3}$), suggesting that dorsal FA is not a mediator or moderator of the relationship between age and phonological activation

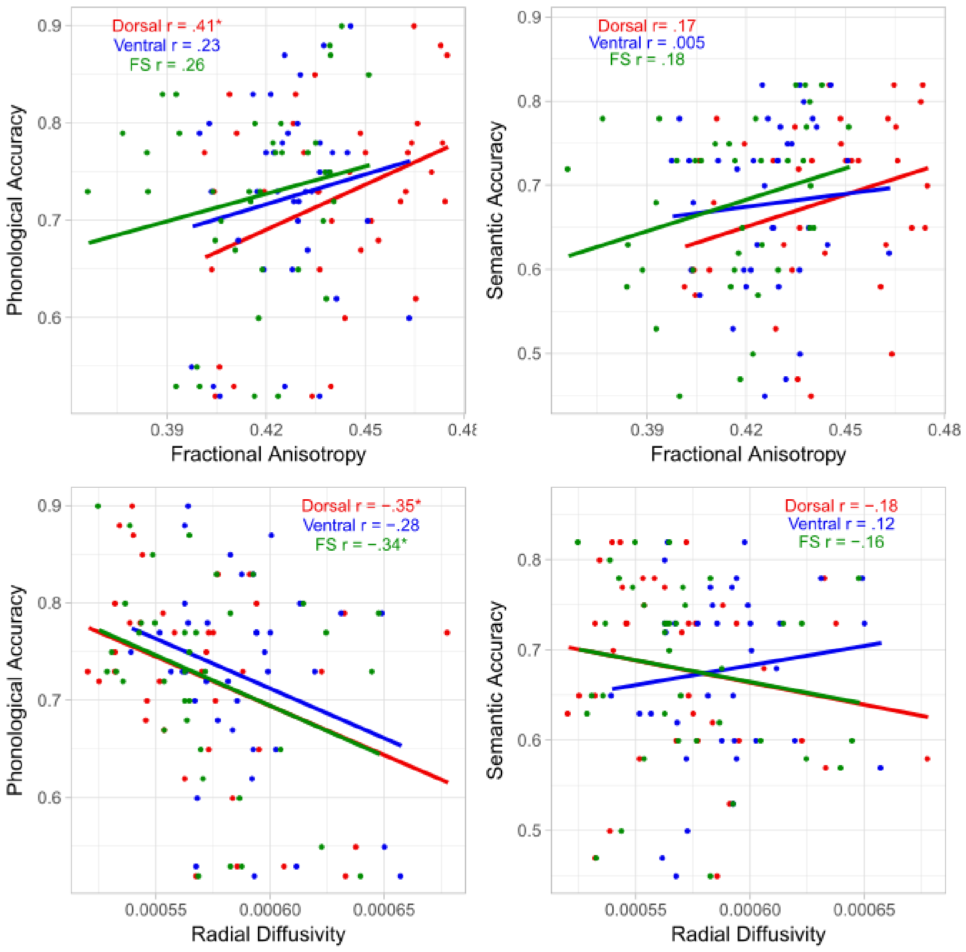

Figure 6. Above are the relationships between accuracy and white matter integrity by condition. Higher FA values are indicative of better white matter while lower RD values are indicative of better white matter. Thus, the direction of the correlation coefficients and regression lines should be opposing.

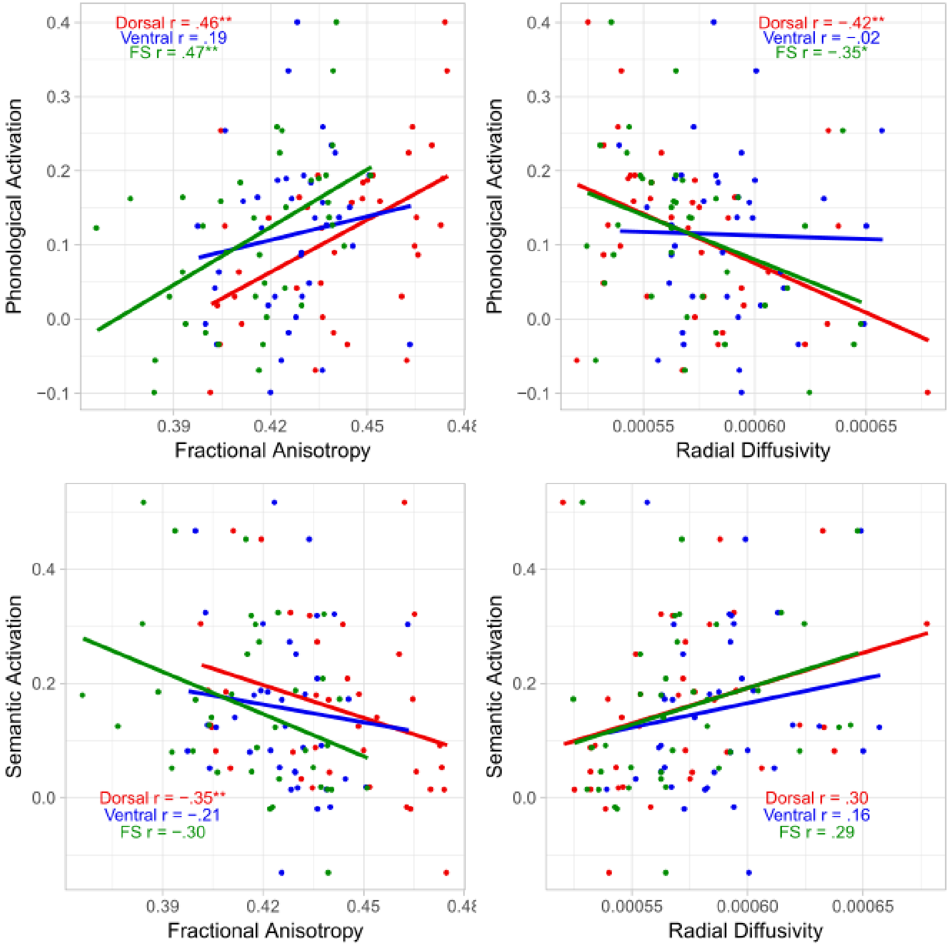

Figure 7. Above are the relationships between activation and white matter integrity by condition. Higher FA values are indicative of better white matter while lower RD values are indicative of better white matter. Thus, the direction of the correlation coefficients and regression lines should be opposing.

Inhibition Deficit Hypothesis Results

Structure-Behavior Relationships

2.1.A Better White matter integrity along the FS tract predicted lower Stroop Effect scores both when white matter was measured by RD (r(36) = .34, *p* = .04) and FA (r(36) = -.34, *p* = .04).

2.1.B Better White matter integrity along the FS tract predicted better phonological accuracy when white matter was measured by RD (r(36) = -.34, $p = .04$) but not when white matter integrity was measured by FA (r(36) = .26, $p = .11$).

2.1.C White matter integrity along the FS tract was not significantly related to semantic accuracy when white matter integrity was measured by RD (r(36) =-.16, p=.34) or FA (r(36) = .18, $p = .28$).

Structure-Activation Relationships

2.2. A Better FS tract integrity predicted greater activation in the phonological ROI when white matter integrity was measured by both RD (r(36) = -.35, $p = .03$) and FA (r(36) = .47, $p = .003$).

2.2.B FS tract integrity was not related to activation in the semantic ROI when white matter integrity was measured by RD (r(36) = .29, $p = .08$) or FA (r(36) = -.30, $p = .06$).

Interaction and Mediation Analyses

2.3.A In a multiple regression model including only FS tract RD and age to predict Stroop Effect, age was a significant predictor ($\beta = .68$, $r_{rdage.stroop} = .55$, $t(35) = 3.87$, $p < .001$) but FS tract RD was not ($\beta = -.10$, $r_{rdstroop.age} = -.10$, $t(35) = -.58$, $p = .57$). Therefore, no interaction was tested. Similarly, in a model including only FS tract FA and age to predict Stroop Effect, age was a significant predictor ($\beta = .73$, $r_{faage.stroop} = .56$, $t(35) =$

3.95, *p* < .001) but FS tract FA was not (β = .17, $r_{\text{fastroop.age}}$ = .15, *t*(35) = .92, *p* = .37). Therefore, no interaction of FS tract FA and age was tested.

2.3.B In a multiple regression model including only FS tract RD and age to predict phonological accuracy, age was not a significant predictor (β = -.35, $r_{\text{rdaccuracy.age}}$ = -.28, *t*(35) = -1.72, *p* = .09) nor was FS tract RD (β = -.12, $r_{\text{faage.accuracy}}$ = -.09, *t*(35) = -.53, *p* = .60). Therefore, no interaction was tested. There was also not a significant bivariate relationship between FS tract FA and phonological accuracy so no interaction was tested.

2.3.C Since there were no significant bivariate relationships between semantic activation white matter integrity along the FS tract, measured by RD or FA, no white matter by age interactions were tested.

2.3.D A mediation analysis using a bootstrap approach with 5000 samples revealed that after controlling for the effect of FS tract RD, the direct effect of age on Stroop Effect remained significant (β= 2.00, *t*(36) = 3.87, *p* < .001). The indirect effect of FS tract RD was insignificant (β = -.19, SE = .30, 95% CI = -.83, .39). This suggests that RD is not a statistically significant mediator or moderator of the age-Stroop Effect relationship. Similarly, after controlling for the effect of FS tract FA, the direct effect of age on Stroop Effect remained significant (β = 2.15, *t*(36) = 3.95, *p* < .001). The indirect effect of FS tract FA was insignificant (β = -.35, SE = .36, 95% CI = -1.13, .32). This suggests that FS tract FA is not a statistically significant mediator or moderator of the age-Stroop Effect relationship.

2.3.E A mediation analysis using a bootstrap approach with 5000 samples revealed that after controlling for the effect of FS tract RD, the direct effect of age on phonological accuracy was no longer significant ($\beta = -1.6\text{x}10^{-3}$, t (36) = -1.73, $p = .09$) despite an insignificant indirect effect of FS RD ($\beta < -.30\text{x}10^{-3}$, SE = $.60\text{x}10^{-3}$, 95% CI = $-1.7\text{x}10^{-3}$, $.80\text{x}10^{-3}$). This suggests that dorsal RD is not a statistically significant mediator but is a significant moderator of the relationship between age and phonological accuracy. Since there was no significant bivariate relationship between phonological accuracy and FS tract FA, no mediation analysis was run with these variables.

2.3.F Since there was no significant bivariate relationship between phonological accuracy and FS tract RD or FA, no mediation analysis was run with these variables.

Chapter 5

Discussion

Transmission Deficit Hypothesis

The Transmission Deficit Hypothesis suggests that age-related differences in language production ability can be explained by the failure of signals to reach their destination; it holds that, with age, signal failures increase across systems of the brain, leading to behavioral declines in the phonological system but not the semantic system. The Transmission Deficit Hypothesis suggests this occurs because in the phonological system the connections are one-to-one but in the semantic system redundant connections compensate for isolated signal transmission failures (MacKay & Burke, 1990). As predicted by this framework, white matter integrity was lower in older relative to younger adults and this lower integrity was uniquely related to behavioral manifestations within the phonological system. Semantic behavior was not predicted by white matter integrity of dorsal or ventral tracts. Moreover, the relationship between dorsal tract white matter integrity and phonological behavior was significantly stronger than the relationship between ventral tracts white matter integrity and semantic behavior.

Analogously, activation in the phonological ROI was significantly predicted by dorsal tract integrity but not ventral tract integrity. While ventral tract integrity did not predict semantic activation, dorsal tract FA did predict activation in the semantic ROI. The strength of the relationship between white matter integrity and activation were statistically equivalent between the phonological and semantic systems. While this result is not predicted by the Transmission Deficit Hypothesis (MacKay & Burke, 1990), it could be explained by the fact that both conditions of the PWI task required speaking. Thus, the semantic condition also may have some degree of phonological task demand which, though observable at the level of activation, may not be strong enough to be observable at the level of behavior. This is consistent with previous literature indicating that different levels of analysis reveal the influence of different biological substrates of cognitive aging (Bennett & Madden, 2014; Madden et al., 2017). This interpretation is also supported by the finding that white matter was neither a moderator nor mediator of age-activation relationships. Additionally, it is important to bear in mind that, because of how the ROIs were selected, younger adults drive effects found in the phonological ROI and older adults drive effects found in the semantic ROI. Because older adults drive semantic activation, the finding that dorsal FA was related to semantic activation could be interpreted in terms of the dedifferentiation hypothesis wherein older adults utilize more diffuse, less specialized brain regions than younger adults (Li et al., 2001).

When controlling for the effect of age, all of the observed structure-behavior and structure-function relationships within the phonological and semantic systems are attenuated to insignificant levels. Though this precludes the possibility of the predicted age by white matter integrity interaction, it does indicate that, in this sample, age and age-related white matter integrity differences are statistically the same thing. One interpretation of this result could be that age is actually acting as a proxy variable for some other biological variable which dually influences cognition and white matter integrity. For example, cardiovascular function (Colcombe, Kramer, McAuley, Erickson, & Scalf, 2004) and diet (McEvoy, Guyer, Langa, & Yaffe, 2017) have been linked to both cognitive performance and white matter integrity (Cermakova et al., 2017; Ong et al., 2018). However, this sample was of healthy older adults who did not have significant health problems. Alternately, the fact that age attenuated structure-behavior and structure-function relationships could indicate that lower white matter integrity itself is the neural instantiation of age. Put in terms of the Transmission Deficit Hypothesis, lower white matter integrity is the basis of transmission deficits which increase with age

(MacKay & Burke, 1990). This interpretation is strengthened by the fact that white matter integrity was a significant moderator of the age- accuracy relationship in the phonological condition. Thus, this study largely supports the predictions and conceptualizations of cognitive aging put forth by the Transmission Deficit Hypothesis.

Inhibition Deficit hypothesis

The predictions of the Inhibition Deficit Hypothesis, which says that lowered inhibition underpins other cognitive deficits (Hasher & Zacks, 1988), were partially supported. First, this study replicated the results of previous research (Liston et al., 2006), finding that white matter integrity along the FS tract integrity predicted inhibition. Next, when RD was used as the metric of white matter integrity, better white matter integrity along the FS also predicted better phonological accuracy. However, when FS tract integrity was measured by FA, phonological accuracy was not related to the white matter integrity of the FS tract. These discrepant findings between measures of white matter might be a result of differences in the sensitivity of the measures, as RD has been shown to be a more sensitive indicator of age-related differences in white matter integrity (Davis et al., 2009).

Counter to the predictions of the Inhibition deficit Hypothesis (Hasher & Zacks, 1988), semantic accuracy was not predicted by FS tract integrity. Results from the structure-activation analysis also indicate that, though better FS tract integrity predicted stronger activation in the phonological ROI, it was unrelated to activation in the semantic ROI. Though this could potentially indicate that the semantic condition of the PWI task does not require inhibition, previous literature indicates that the semantic condition causes interference effects (Lupker, 1979) and requires more inhibitory control. One interpretation of these opposing findings between the phonological and semantic systems could be that communication with inhibition-related centers, being just one factor among many impacting performance on the PWI task, simply does not exert a statistically significant effect when compared to the influence of other cognitive processes. Said another way, these results—both behavioral and functional—support the notion that multiple cognitive processes are required for speaking and that FS tract integrity alone is not enough to explain speaking ability.

The conclusion that inhibition is a marginal factor in driving age-related language differences is further supported by interaction and mediation analyses. When accounting for age, the bivariate relationships between FS tract integrity and behavior or activation were attenuated to insignificant levels, again suggesting that age and white matter integrity account for the same variance. FS tract integrity was also not a statistical mediator of any of the age-behavior relationships, precluding this study from providing firm evidence that white matter integrity underpins inhibition driven differences in cognition between older and younger adults. FS tract RD, however, was a moderator of the age-phonological accuracy relationship. This indicates that a relationship may exist between FS tract integrity and the ability to utilize inhibition resources for language production, but this relationship is secondary to the influence of other factors.

Neural Aging

Though this study found no significant relationships between activation and behavior, there are still important implications for neural aging theories. First, this study does confirm that the relationships between white matter integrity and behavior are not exclusive to phonological, semantic, or inhibition systems. Rather, age-related differences in white matter integrity occur across multiple tracts. Further, white matter integrity across various tracts statistically accounts for age-related behavioral differences in both the language and inhibition domains. This adds to existing evidence that white matter integrity may be a neural mechanism underpinning cognitive aging (Bennett & Madden, 2014; Head et al., 2004). Second, when age was included in multiple regression models, all significant bivariate structure-behavior and structure-activation relationships were attenuated to nothing, indicating that age and white matter integrity are accounting for the same variance. The finding that age attenuated the bivariate relationships between white matter integrity and behavior as well as between white matter and activation suggests that differences in white matter integrity may underlie other processes of neural reorganization observed with aging.

In particular this study provides evidence for the CRUNCH framework over the dedifferentiation or compensation accounts. The dedifferentiation account of neural aging would suggest that older adults process information using more diffuse networks than younger adults and, as a result, behavioral outcomes would be worse in older adults (Li et al., 2001). While Rizio and colleagues did find that older adults in this sample engaged in more diffuse processing in the semantic condition and were poorer performers (Rizio et al., 2017), the current study found no age by white matter interactions. That is, older adults rely on the same white matter networks as younger adults. Compensation accounts of neural aging predict that older adults engage in additional activation to maintain behavioral performance in spite of less efficient processing (Cabeza, 2002) but, in this study, behavioral performance was not maintained in older adults. The CRUNCH framework, which represents a task-difficulty indexed hybrid of the differentiation and compensation models (Reuter-Lorenz & Cappell, 2008), suggests that processing becomes less efficient with age and that additional processing is needed to maintain behavioral performance but, during very difficult tasks, the additional processing becomes insufficient. In this study, older and younger adults utilized the same white matter networks but younger adults outperformed older adults, therefore, this study could be interpreted within the CRUNCH model of neural aging.

Limitations, Strengths, Future Directions, & Conclusion

While this cross-sectional study is inherently limited its inability to assess the directionality of structure-behavior or structure-function relationships, it nonetheless provides novel information regarding how white matter integrity relates to older adults' language ability. The cross-sectional nature of this study also means no comment can be made on the underlying mechanisms of change at work during aging; only longitudinal research can speak to the actual processes of change occurring across time. Thus, it is critical that future research utilize intervention-based and longitudinal methods to probe both structure-behavior and structure-function relationships. Emergent literature on brain connectivity has also indicated that functional connectivity, more so than structural connectivity, may drive certain aspects of cognition—particularly executive function (Madden et al., 2017). As this study only utilizes structural metrics to quantify connectivity, only some types of connectivity-behavior relationships are captured. Going forward, it will be critical for research to integrate findings from multiple types of connectivity analysis, to gain a full picture of cognitive aging.

Another limitation of this study is the number of comparisons made to fully address the research questions asked. Making multiple comparisons—among other factors such as small sample sizes—can lead to incorrect conclusions (Munafò et al., 2017). To address this, I used post-hoc corrections and power analyses (see Appendix B). Six comparisons were made to test for differences in white matter integrity between age groups and six novel comparisons were made between white matter indices (dorsal, ventral, and FS integrity measured by FA and RD) and each behavioral or activation outcome. Though cumulatively this resulted in 36 comparisons

(6 for each phonological accuracy, phonological activation, semantic accuracy, semantic activation, Stroop effect, and age differences), each set of six comparisons comprises a separate family of predictions. Bonforroni correction was used in a post-hoc correction (see Appendix B) to test for potential type 1 errors. At the corrected significance value of .008, all of the observed between-group differences hold as do relationships between white matter integrity and phonological activation. However, none of the significant behavioral effects survived this correction. Though this post-hoc analysis indicates that results should be interpreted with caution, Bonferroni corrections are widely considered overly conservative. Given the convergence this study has with existing literature, I expect the conclusion drawn from this study hold. Moreover, post-hoc power analyses (see Appendix B) indicate that if anything this study was underpowered to detect significant effects should they exist and yet this study did identify effects. In addition, precautions were taken in the methods to limit the number of comparisons made and minimize researcher bias.

Finally, it is important to consider the role third variables could play in driving these results. Health factors, which may be systematically different between younger and older adults (König et al., 2018; Levy & Myers, 2004; Vaccaro & Huffman, 2017), are related to both white matter integrity (Cermakova et al., 2017; Ong et al., 2018) and cognition (Colcombe et al., 2004; McEvoy et al., 2017). However, given previous evidence that white matter integrity is lower in older adults (Bennett & Madden, 2014; Head et al., 2004) and the fact that this study utilized a sample of older adults who were healthy, it is unlikely that the structure-behavior and structure-function relationships evident in this study were actually differences in health. Even so, an important future direction of this research line would be to investigate which factors, such as health behaviors, might attenuate structure-behavior and structure-activation relationships. Given that this work largely supports the Transmission Deficit Hypothesis (MacKay & Burke, 1990), future research should also explore how naturally occurring knowledge structure differences can be leveraged to alleviate the behavioral and functional consequences of lower white matter integrity. In sum, this study provides strong evidence that white matter integrity is linked to language production ability across ages and further suggests that lower white matter integrity associated with aging may be a biological substrate of age-related differences in language ability.

Appendix A

Tractography Procedure

	Control Tracts	**Semantic Tracts**		**Phonological Tracts**	
	Frontal Striatal Tract (FS)	**Inferior Longitudinal Fasiculus (ILF)**	**Middle Longitudinal Fasiculus (MDLF)**	**Frontal Aslant Tract (FAT)**	**Superior Longitudinal Fasiculus (SLF) III /Arcuate Fasiculus (AF)**
Seed	Dorsolateral Prefrontal Cortex as defined as a 8mm sphere centered at X=22 Y=87 Z=38	Temporal pole as defined by the Harvard Oxford Cortical Atlas thresholded at 25%	Angular Gyrus as defined by the Harvard Oxford Cortical Atlas thresholded at 25%	Supplementary motor cortex (SMA) and pre-SMA as defined by the Harvard Oxford Cortical Atlas thresholded at 25%	Inferior Frontal Gyrus pars operculus, as defined by Harvard Oxford Cortical Atlas thresholded at 25%
Target	Caudate defined by the Harvard Oxford Atlas thresholded at 25%,	Temporo- occipital portion of the Inferior Frontal gyrus defined by the Harvard Oxford Cortical Atlas thresholded at 25%	Temporal Pole defined by the Harvard Oxford Cortical Atlas thresholded at 25%	Brocas area defined by the Harvard Oxford Cortical atlas Inferior Frontal Gyrus pars operculus and pars triangulars thresholded at 25%	Superior Temporal Gyrus defined by Harvard Oxford Cortical Atlas thresholded at 25%

	Control Tracts	Semantic Tracts		Phonological Tracts	
	Frontal Striatal Tract (FS)	**Inferior Longitudinal Fasiculus (ILF)**	**Middle Longitudinal Fasiculus (MDLF)**	**Frontal Aslant Tract (FAT)**	**Superior Longitudinal Fasiculus (SLF) III /Arcuate Fasiculus (AF)**
Waypoints	None	Targets were included as waypoints	Targets and posterior Superior Temporal Gyrus defined by the Harvard Oxford Cortical Atlas thresholded at 25%	Target were included as waypoints	White matter seed located at MNI coordinates X=59-70 Y=38-48, Z=44-53
Exclusion Masks	Hemisphere mask, superior boundary mask (all voxels above z=51), a posterior boundary mask (all voxels posterior to y=49), and the cerebellum and brainstem defined by the Harvard Oxford Atlas thresholded at 25%,	Midline hemisphere mask, a box at x=60-68 y=25-50 z=38, a box at x=68-78 y=45-57, z=36-41 and the left putamen, insular cortex and anterior part of the superior temporal gyrus defined by the Harvard Oxford Atlas thresholded at 5%,	Midline hemisphere mask, a sagittal slice mask at X=58, a posterior boundary at Y=31, as well as the putamen, insula, and posterior and anterior portions of the Middle Temporal Gyrus defined by the Harvard Oxford Atlas thresholded at 25%, and a box at X=58-74, Y=56-61, Z=39-53	Frontal pole, middle frontal gyrus and precentral gyrus defined by the Harvard Oxford Atlas thresholded at 25%, a hemisphere midline, a line at Y=48, and a box covering the ventricles at X=27-62, Y=48-74 and Z=23-43	putamen, insula , and posterior middle temporal gyrus defined by the Harvard Oxford Atla, thresholded at 25% and a hemisphere midline

	Control Tracts	**Semantic Tracts**		**Phonological Tracts**	
	Frontal Striatal Tract (FS)	**Inferior Longitudinal Fasiculus (ILF)**	**Middle Longitudinal Fasiculus (MDLF)**	**Frontal Aslant Tract (FAT)**	**Superior Longitudinal Fasiculus (SLF) III /Arcuate Fasiculus (AF)**
Thresholding	1% - Tracks could not be reliably tracked using any higher threshold. This indicates that 99% of streamlines were kept	10%- This indicates that 90% of streamlines were kept.			
Combining	Tractography was performed from seed to target and from target to seed. These directions were combined after thresholding and only common streamlines were kept. However for a minority of cases in the ILF, one direction was not reliably tracking. In this case the other direction was still kept.				
Finalizing	In the finalization stage, individual-level tracts were masked with the individuals full FA-map to mask-out any spurious non-white matter. A threshold of 25% was used/ This indicates that the likelihood that each voxel in the final tract is actually white matter in that individual is 75%				

Appendix B

Post-Hoc Analyses

For the between group comparisons of white matter, the smallest observed effect was the between group differences in ventral RD (F = 8.00). A post-hoc power analysis revealed that this study observed 15% power and as such was underpowered to properly detect differences in the two groups. Nonetheless, group differences were detected. Moreover, since six between groups comparisons were made a Bonforroni correction can be used to test for potential type 1 errors; at the corrected significance value of $p = .008$, the observed between-group differences hold.

Post hoc power analysis indicated that this study observed 7% power to detect significant bivariate correlations between white matter and behavior (semantic accuracy, phonological accuracy, and Stroop effect) at the 95% confidence level, when effect size was indexed by the smallest observed white matter-behavior relationship (r(35) = .005). Thus it is possible that this study was underpowered to detect relationships, particularly between semantic accuracy and white matter indices where effect sizes were smallest. Even so, the fact that these effect sizes are smaller than those observed when measuring the relationship between white matter and semantic accuracy or Stroop Effect, indicates support for the conclusions drawn.

In total, six novel comparisons were made between white matter indices (dorsal, ventral, and FS integrity measured by FA and RD and each behavioral or functional outcome. Though cumulatively this resulted in 30 comparisons (6 for each phonological accuracy, phonological activation, semantic accuracy, semantic activation, Stroop effect), I will argue that each set of six comparisons comprises a separate family. As such a Bonferroni correction was calculated (.05 / 6 =.008). The significant relationships between white matter integrity and phonological activation did survive this correction. However none of the significant behavioral effects survived this correction so results should be interpreted with caution. Even so, Bonferroni corrections are considered too strict by field standard and, given the convergence this study has with existing literature, I expect these conclusions should still hold.

Works Cited

Bach, M. (1996). The Freiburg Visual Acuity test--automatic measurement of visual acuity. *Optom Vis Sci, 73*(1), 49-53.

Bennett, I. J., & Madden, D. J. (2014). Disconnected aging: cerebral white matter integrity and age-related differences in cognition. *Neuroscience, 276*, 187-205.

Britt, A. E., Ferrara, C., & Mirman, D. (2016). Distinct Effects of Lexical and Semantic Competition during Picture Naming in Younger Adults, Older Adults, and People with Aphasia. *Front Psychol, 7*, 813. doi:10.3389/fpsyg.2016.00813

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) phase II: 930 new normative photos. *PLoS One, 9*(9), e106953. doi:10.1371/journal.pone.0106953

Burke, D. M. (1997). Language, aging, and inhibitory deficits: evaluation of a theory. *J Gerontol B Psychol Sci Soc Sci, 52*(6), P254-264.

Burke, D. M., & Mackay, D. G. (1997). Memory, language, and ageing. *Philos Trans R Soc Lond B Biol Sci, 352*(1363), 1845-1856. doi:10.1098/rstb.1997.0170

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? In *Journal of memory and language* (Vol. 30, pp. 542-579): Elsevier.

Cabeza, R. (2002). Hemispheric asymmetry reduction in older adults: the HAROLD model. *Psychol Aging, 17*(1), 85-100.

Cermakova, P., Muller, M., Armstrong, A. C., Religa, D., Bryan, R. N., Lima, J. A. C., & Launer, L. J. (2017). Subclinical Cardiac Dysfunction and Brain Health in Midlife:

CARDIA (Coronary Artery Risk Development in Young Adults) Brain Magnetic Resonance Imaging Substudy. *J Am Heart Assoc, 6*(12). doi:10.1161/JAHA.117.006750

Chen, Q., & Mirman, D. (2012). Competition and cooperation among similar representations: toward a unified account of facilitative and inhibitory effects of lexical neighbors. *Psychol Rev, 119*(2), 417-430. doi:10.1037/a0027175

Colcombe, S. J., Kramer, A. F., McAuley, E., Erickson, K. I., & Scalf, P. (2004). Neurocognitive aging and cardiovascular fitness. *Journal of Molecular Neuroscience, 24*(1), 9-14.

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Hum Brain Mapp, 8*(2-3), 109-114.

Davis, S. W., Dennis, N. A., Buchler, N. G., White, L. E., Madden, D. J., & Cabeza, R. (2009). Assessing the effects of age on long white matter tracts using diffusion tensor tractography. *Neuroimage, 46*(2), 530-541.

Davis, S. W., Dennis, N. A., Daselaar, S. M., Fleck, M. S., & Cabeza, R. (2008). Que PASA? The posterior-anterior shift in aging. *Cereb Cortex, 18*(5), 1201-1209. doi:10.1093/cercor/bhm155

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition, 42*(1-3), 287-314.

Diaz, M. T., Johnson, M. A., Burke, D. M., & Madden, D. J. (2014). Age-related differences in the neural bases of phonological and semantic processes. *J Cogn Neurosci, 26*(12), 2798-2811. doi:10.1162/jocn_a_00665

Dick, A. S., Bernal, B., & Tremblay, P. (2014). The language connectome: new pathways, new concepts. *Neuroscientist, 20*(5), 453-467. doi:10.1177/1073858413513502

Fjell, A. M., Sneve, M. H., Grydeland, H., Storsve, A. B., Amlien, I. K., Yendiki, A., & Walhovd, K. B. (2017). Relationship between structural and functional connectivity change across the adult lifespan: A longitudinal investigation. *Hum Brain Mapp, 38*(1), 561-573. doi:10.1002/hbm.23403

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. *J Psychiatr Res, 12*(3), 189-198.

Gazes, Y., Bowman, F. D., Razlighi, Q. R., O'Shea, D., Stern, Y., & Habeck, C. (2016). White matter tract covariance patterns predict age-declining cognitive abilities. *Neuroimage, 125*, 53-60. doi:10.1016/j.neuroimage.2015.10.016

Guasch, M., Boada, R., Ferré, P., & Sánchez-Casas, R. (2013). NIM: A Web-based Swiss army knife to select stimuli for psycholinguistic studies. *Behav Res Methods, 45*(3), 765-771. doi:10.3758/s13428-012-0296-8

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *Psychology of learning and motivation, 22*, 193-225.

Hayes, A. F. (2016). The PROCESS macro for SPSS and SAS. *Retrieved from*.

Head, D., Buckner, R. L., Shimony, J. S., Williams, L. E., Akbudak, E., Conturo, T. E., . . . Snyder, A. Z. (2004). Differential vulnerability of anterior white matter in nondemented aging with minimal acceleration in dementia of the Alzheimer type: evidence from diffusion tensor imaging. *Cerebral cortex, 14*(4), 410-423.

Hedden, T., Schultz, A. P., Rieckmann, A., Mormino, E. C., Johnson, K. A., Sperling, R. A., & Buckner, R. L. (2016). Multiple Brain Markers are Linked to Age-Related Variation in Cognition. *Cereb Cortex, 26*(4), 1388-1400. doi:10.1093/cercor/bhu238

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience, 8*(5), 393-402.

James, L. E. (2004). Meeting Mr. Farmer versus meeting a farmer: specific effects of aging on learning proper names. *Psychol Aging, 19*(3), 515-522. doi:10.1037/0882-7974.19.3.515

Jones, D. K., Knösche, T. R., & Turner, R. (2013). White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. *Neuroimage, 73*, 239-254. doi:10.1016/j.neuroimage.2012.06.081

König, M., Drewelies, J., Norman, K., Spira, D., Buchmann, N., Hülür, G., . . . Demuth, I. (2018). Historical trends in modifiable indicators of cardiovascular health and self-rated health among older adults: Cohort differences over 20 years between the Berlin Aging Study (BASE) and the Berlin Aging Study II (BASE-II). *PLoS One, 13*(1), e0191699. doi:10.1371/journal.pone.0191699

Laver, G. D., & Burke, D. M. (1993). Why do semantic priming effects increase in old age? A meta-analysis. *Psychol Aging, 8*(1), 34-43.

Levy, B. R., & Myers, L. M. (2004). Preventive health behaviors influenced by self-perceptions of aging. *Prev Med, 39*(3), 625-629. doi:10.1016/j.ypmed.2004.02.029

Li, S. C., Lindenberger, U., & Sikström, S. (2001). Aging cognition: from neuromodulation to representation. *Trends Cogn Sci, 5*(11), 479-486.

Liston, C., Watts, R., Tottenham, N., Davidson, M. C., Niogi, S., Ulug, A. M., & Casey, B. J. (2006). Frontostriatal microstructure modulates efficient recruitment of cognitive control. *Cereb Cortex, 16*(4), 553-560. doi:10.1093/cercor/bhj003

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: the neighborhood activation model. *Ear Hear, 19*(1), 1-36.

Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory & Cognition, 7*(6), 485-495.

Lutkenhoff, E. S., Rosenberg, M., Chiang, J., Zhang, K., Pickard, J. D., Owen, A. M., & Monti, M. M. (2014). Optimized brain extraction for pathological brains (optiBET). *PLoS One, 9*(12), e115551. doi:10.1371/journal.pone.0115551

MacKay, D. G., & Burke, D. M. (1990). Chapter five cognition and aging: a theory of new learning and the use of old connections. In T. M. Hess (Ed.), *Aging and Cognition: Knowledge Organization and Utilization* (pp. 213-263): Elsevier Science Publishers B.V. (North-Holland).

Madden, D. J., Parks, E. L., Tallman, C. W., Boylan, M. A., Hoagey, D. A., Cocjin, S. B., . . . Potter, G. G. (2017). Sources of disconnection in neurocognitive aging: cerebral white-matter integrity, resting-state functional connectivity, and white-matter hyperintensity volume. *Neurobiology of aging, 54*, 199-213.

May, C. P., Zacks, R. T., Hasher, L., & Multhaup, K. S. (1999). Inhibition in the processing of garden-path sentences. *Psychol Aging, 14*(2), 304-313.

McEvoy, C. T., Guyer, H., Langa, K. M., & Yaffe, K. (2017). Neuroprotective Diets Are Associated with Better Cognitive Function: The Health and Retirement Study. *J Am Geriatr Soc, 65*(8), 1857-1862. doi:10.1111/jgs.14922

Moreno-Martínez, F. J., & Montoro, P. R. (2012). An ecological alternative to Snodgrass & Vanderwart: 360 high quality colour images with norms for seven psycholinguistic variables. *PLoS One, 7*(5), e37527. doi:10.1371/journal.pone.0037527

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., du Sert, N. P., . . . Ioannidis, J. P. (2017). A manifesto for reproducible science. *Nature Human Behaviour, 1*, 0021.

Ong, I. M., Gonzalez, J. G., McIlwain, S. J., Sawin, E. A., Schoen, A. J., Adluru, N., . . . Yu, J. J. (2018). Gut microbiome populations are associated with structure-specific changes in white matter architecture. *Transl Psychiatry, 8*(1), 6. doi:10.1038/s41398-017-0022-5

Pfefferbaum, A., Adalsteinsson, E., & Sullivan, E. V. (2005). Frontal circuitry degradation marks healthy adult aging: evidence from diffusion tensor imaging. *Neuroimage, 26*(3), 891-899.

Redfern, M. S., Jennings, J. R., Martin, C., & Furman, J. M. (2001). Attention influences sensory integration for postural control in older adults. *Gait Posture, 14*(3), 211-216.

Reuter-Lorenz, P. A., & Cappell, K. A. (2008). Neurocognitive aging and the compensation hypothesis. *17(3)*, 177-182.

Rizio, A. A., & Diaz, M. T. (2016). Language, aging, and cognition: frontal aslant tract and superior longitudinal fasciculus contribute toward working memory performance in older adults. *Neuroreport, 27*(9), 689-693. doi:10.1097/WNR.0000000000000597

Rizio, A. A., Moyer, K. J., & Diaz, M. T. (2017). Neural evidence for phonologically based language production deficits in older adults: An fMRI investigation of age-related differences in picture-word interference. *Brain and behavior, 7*(4).

Shafto, M. A., & Tyler, L. K. (2014). Language in the aging brain: the network dynamics of cognitive decline and preservation. *Science, 346*(6209), 583-587. doi:10.1126/science.1254404

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., . . . Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage, 23 Suppl 1*, S208-219. doi:10.1016/j.neuroimage.2004.07.051

Stamatakis, E. A., Shafto, M. A., Williams, G., Tam, P., & Tyler, L. K. (2011). White matter changes and word finding failures with increasing age. *PLoS One, 6*(1), e14496. doi:10.1371/journal.pone.0014496

Taylor, J. K., & Burke, D. M. (2002). Asymmetric aging effects on semantic and phonological processes: naming in the picture-word interference task. *Psychol Aging, 17*(4), 662-676.

Vaccaro, J. A., & Huffman, F. G. (2017). Dietary Risk Factors by Race/Ethnicity, Age-Group, and Gender in a Representative Sample of US Older Adults. *J Nutr Health Aging, 21*(6), 704-709. doi:10.1007/s12603-016-0744-8

Vitevitch, M. S., & Luce, P. A. (2016). Phonological neighborhood effects in spoken word perception and production. *Annual Review of Linguistics, 2*, 75-94.

Zacks, R., & Hasher, L. (1997). Cognitive gerontology and attentional inhibition: a reply to Burke and McDowd. *J Gerontol B Psychol Sci Soc Sci, 52*(6), P274-283.