DISCOURSE PROCESSING ABILITIES IN AGEING: INFLUENCE OF WORKING MEMORY CAPACITY ON REFERENCE RESOLUTION

A thesis submitted as fulfilment of the requirements for the degree
of Doctor of Philosophy
at University of Canterbury

by Maryam Ghaleh

July 2014

DECLARATION BY AUTHOR

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Maryam Ghaleh

LIST OF PRESENTATIONS BY THE AUTHOR RELEVANT TO THE THESIS

1. Ghaleh, M., McAuliffe, M., Moran, C., Neumann, E. and Lashell, P. (2013). *Effects of age and working memory on the comprehension of anaphora.* Chicago, IL, USA: American Speech-Language-Hearing Association (ASHA) Annual Convention, 14-16 Nov 2013.
2. Ghaleh, M., McAuliffe, M., Moran, C., Neumann, E. and Lashell, P. (2014, accepted). *Cognitive processing underlying anaphora resolution: effects of individual differences in working memory capacity.* Orlando, Florida, USA: American Speech-Language-Hearing Association (ASHA) Annual Convention, 20-22 Nov 2014.

ACKNOWLEDGEMENTS

At the beginning, PhD journey seemed like stepping on an endless foggy path. One cloudy morning, I started the journey passionately with a blurry image of the route. Unseen turns, rocky roads, rains, snows and hailstorms could have been exhausting without the help, advice and the encouragements I received from the most amazing people in my life. I owe the journey to those who showed me the path, who trusted me on completing the journey and helped me on the unfortunate days. I will not be able to express my gratitude in words to all those people who helped me. However, I would like to express my deep gratitude…

To Associate Professor Megan McAuliffe, who first accepted me as her Ph.D. student and trusted me on this journey. Without her help and encouragement, I could not start and complete this thesis. I would like to thank her for her patience, perfect supervision, valuable advices, useful comments and endless support, and for preparing me for my next journey. She taught me how to critically think, read, and write, and how to become an independent researcher. Megan, I have learned a lot from you - much more than I had expected to achieve during my PhD - and I can never thank you enough for that.

To Associate Professor Catherine Moran, and Dr Ewald Neumann, for their great supervisions, encouragements, valuable advice, and ongoing supports. Their conversations and advices enlightened my way of thinking. Without their guidance and thought-provoking ideas, I could have been lost early on the way. Catherine and Ewald, thanks for helping me see the path more clearly to make my way through the fog.

To Dr Patrick Lashell for being one of the greatest teachers I have ever had, and for his generous help. His great knowledge and smart ideas have been very helpful and inspiring.

He taught me the essentials of research and made me interested in statistics and programming. Pat, I owe all my technical skills to you.

To my family, to whom I owe everything. Without their support, I could not start, continue and complete my PhD in a paradise far away from home. Thanks for your endless encouragement and caring. Thanks for patiently listening to all my complaints on my hard days and for giving me strengths to move on.

To all my participants and those who helped me with recruiting participants for this study. My special thanks to Russell and the other members of Operation Friendship.

To Professor Nader Jahangiri and Associate Professor Mahmoud Reza Azarpazhooh for introducing me to the excitement of research, for their encouragements and for supporting me in getting into this PhD program.

To all my friends in Christchurch whose presence made the journey much more enjoyable. My endless thanks to Farzaneh and Shahrzad for proofreading this thesis.

Finally yet importantly, to the New Zealand Institute of Language, Brain and Behaviour (NZILBB) and the Canterbury Medical Research Foundation (CMRF) for their support. I am deeply grateful to these institutes for providing me with a joint scholarship, which enabled me to pursue my PhD.

Maryam Ghaleh

ABSTRACT

Maintaining health and quality of life into old age is a critical issue facing society today. Language, and in particular language comprehension, is vulnerable to the processes of ageing (Au, Albert, & Obler, 1989; Kynette & Kemper, 1986; Nicholas, Obler, Albert, & Goodglass, 1985; Shewan & Henderson, 1988). An improved understanding of language processing and ageing will assist in distinguishing language difficulties in normal ageing from those in pathological ageing and aphasia (Maxim & Bryan, 1994) and, potentially, optimises communication throughout life. The current thesis focuses on a specific component of language comprehension—anaphora resolution[1].

Anaphora resolution occurs frequently in everyday discourse and has been reported to decline with ageing (Cohen, 1979; Light & Capps, 1986; Ulatowska, Hayashi, Cannito, & Fleming, 1986). This thesis explored anaphora resolution relative to two key variables: ageing and working memory. Ageing was chosen as a variable as anaphora resolution has been shown to be affected by age (Cohen, 1979; Light & Capps, 1986; Ulatowska et al., 1986). Working memory was chosen as working memory is thought to underlie key aspects of discourse comprehension such as building a mental structure of discourse and updating the information (Brébion, 2003; Hasher & Zacks, 1988; Radvansky, Copeland, & Hippel, 2010; Radvansky, Lynchard, & von Hippel, 2009).

[1] Anaphora resolution is the process of finding what/whom an anaphora refers to in a discourse. Anaphora is the most common type of reference in English that is used to refer to discourse entities that have been previously mentioned in the discourse. For example, in the sentence *I bought a sandwich and ate it immediately*, *it* is an anaphoric pronoun referring back to *sandwich*. *Sandwich* is the referent of the anaphora. Further information about anaphoric references and anaphora resolution can be found on page 28.

Anaphora resolution was investigated using two key paradigms. The first focussed on anaphora resolution in a reading comprehension task. Performance was assessed using accuracy of response. The second employed Gernsbacher's (1989) probe-response paradigm. The probe- response paradigm allowed examination of specific working memory processes underlying discourse comprehension, namely; a) storing and maintaining information in working memory (i.e., laying the foundation of the discourse structure); and b) updating information stored in working memory through suppressing the irrelevant discourse information. Storage and maintenance of the information was assessed by examining whether participants utilised "advantage of first mention" (Gernsbacher, 1990). Suppression[2] was evaluated by investigating whether the accessibility of nonreferent names decreased in participants' working memory after they read anaphoric pronouns in sentences.

This approach aimed to answer the following questions: 1) Do age and working memory capacity affect anaphora resolution in a comprehension task?; 2) Do age and working memory affect advantage of first mention in a probe recognition task?; and 3) Does age affect suppression of irrelevant information in an anaphora resolution task? In Chapter 3, Gernsbacher's (1989) original study was replicated. In Chapter 4 the same questions were examined, with the addition of a higher working memory load.

For both studies, 30 younger and 30 older participants completed two comprehension experiments followed by an assessment of working memory capacity (reading span task). The comprehension experiments each contained a reading comprehension task and a probe recognition task. The reading comprehension task introduced two discourse characters (either a male or female name), one of which was referred to later in the text, using an anaphoric

[2] For the purposes of this thesis, *suppression* is defined as the process of decreasing the activation of already-activated information in working memory. Further discussion of the concept of suppression can be found on page 9.

pronoun. Comprehension questions always asked about the referents of the anaphoric pronouns. Participants' accuracy in answering each comprehension question was indicative of their ability to resolve anaphora. Response times in the recognition task provided measures of the accessibility of: a) first and second mentioned names, and b) referent and nonreferent names.

Chapters 3 and 4 found that, regardless of the tasks' working memory storage demands, older adults were less accurate than younger adults in the comprehension of anaphoric pronouns. Comprehension accuracy was related to working memory capacity, such that individuals with higher working memory capacity exhibited higher accuracy of response in the comprehension task. In addition, working memory capacity affected the accessibility of first and second mentioned names in the discourse suggesting that working memory capacity might influence the process of laying the foundation for the mental representation of comprehension. An ageing effect was observed on the suppression process during anaphora resolution under high working memory load only. When working memory load was low, neither younger nor older participants suppressed the accessibility of the nonreferents by the time they finished reading the sentences. This suggested that anaphora resolution might be postponed in less demanding tasks. However, under higher working memory load, younger adults, but not older adults, suppressed the accessibility of the nonreferents by the time they finished reading the sentence. It was therefore suggested that age-related changes in anaphora resolution abilities might be mediated by a decline in inhibitory functions [3] that are responsible for suppressing the already-activated information that are no longer relevant to the task goals.

[3] *Inhibitory abilities* in this thesis is used as a general term that refers to different executive functions responsible for inhibiting a response or information by preventing it from becoming activated, or suppression of already-activated information.

The final study of the thesis (Chapter 5) aimed to determine why younger adults delayed the process of anaphora resolution in Experiment 1 (See Chapter 3), but completed the process by the time they finished reading the sentences in Experiment 2 (See Chapter 4). Specific questions addressed were: 1) Was comprehension accuracy affected by working memory storage load and the syntactic structure of the sentences?; 2) Do younger adults suppress the accessibility of the nonreferents by the time they reach the end of the sentence, in simpler sentences with increased storage load and late disambiguation?; and, 3) Do younger adults suppress the accessibility of nonreferents by the time they reach the end of the sentence, in more syntactically complex sentences with low storage load and prior disambiguation?. Forty younger participants completed four separate comprehension experimental tasks followed by a reading span test. A similar experimental approach was employed to that described in Chapters 3 and 4; however working memory storage load, syntactic complexity, and time-course for providing contextual information were manipulated.

Results of Chapter 5 found that participants' accuracy declined in more syntactically complex sentences. A decline in accuracy appeared indicative of the tasks' higher processing demands and demonstrated that prior disambiguation was not facilitating the resolution of anaphora. Results from the recognition task showed that in sentences of increased syntactic complexity, participants suppressed the accessibility of nonreferents by the time they finished reading the sentence. It was suggested that higher processing demands of syntactically complex sentences, rather than a facilitating effect of earlier disambiguation in these sentences, contributed to the earlier suppression of nonreferents.

In summary, this thesis demonstrated that older adults were less accurate than younger adults in comprehending anaphoric pronouns. Moreover, working memory capacity

positively influenced comprehension accuracy and affected the advantage of first mention of discourse entities. It was suggested that individual differences in working memory capacity might affect the ability to lay foundations for discourse comprehension. Furthermore, older adults showed no suppression of nonreferents during processing of anaphora, regardless of working memory storage load. It appears possible that older adults' difficulty in anaphora resolution might be due to an inability to suppress irrelevant discourse information. Findings from the present study suggest that ageing may negatively affect the comprehension of linguistic structures for which more than one meaning could be inferred. While further exploration of this finding is required, it is possible that communication strategies could be devised to minimise the use of structures with more than one meaning—with the aim of improving and maintaining communication in older adults. Ultimately, determining the underlying causes of language impairments in both healthy ageing and neurological disease will help to improve speech-language therapy methods for these populations.

KEY WORDS

Ageing, Discourse Processing, Anaphora Resolution, Structure Building Framework, Advantage of First Mention, Working Memory Capacity, Suppression, Inhibitory Decline.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. CHAPTER ONE

A Review of the Literature

1.1 INTRODUCTION

As people age, changes are observed in both physiological and cognitive functions. Language comprehension is one cognitive ability reported to decline with ageing (Au et al., 1989; Cohen, 1979; Daneman & Carpenter, 1980; Hancock, Fisk, & Rogers, 2001; Kynette & Kemper, 1986; Nicholas et al., 1985; Shewan & Henderson, 1988; Ulatowska et al., 1986). Changes in language comprehension have the potential to affect older adults' quality of life, including their social lives (Bergland & Narum, 2007; Hancock et al., 2001). Despite efforts devoted to the investigation of age-related changes in comprehension abilities, there is still no agreement on the nature of these deficits. A better understanding of the nature of comprehension difficulties faced by older adults is required to improve communication with them and to develop more efficient methods of speech-language therapy for this population. Furthermore, a comprehensive knowledge of the age-related changes in language processing is required prior to investigating language difficulties caused by strokes and neurodegenerative diseases, as these mostly occur in older adults (Maxim & Bryan, 1994)

Older adults commonly exhibit minimal difficulty comprehending words (Federmeier, van Petten, Schwartz, & Kutas, 2003; Giaquinto, Ranghi, & Butler, 2007) and individual sentences (Wingfield, Peelle, & Grossman, 2003). However, age-related changes have been observed in the comprehension of discourse, particularly above the surface level of processing—that is, in tasks that require inference to be made based on the integration of text information and previous knowledge. Studies have reported that older adults exhibit significant difficulty with the comprehension of implicit discourse (Borella, Ghisletta, & De Ribaupierre, 2011; Cohen, 1979; Hancock et al., 2001; Hannon & Daneman, 2009; Ulatowska et al., 1986) and when inferring potentially ambiguous discourse concepts such as metaphors, stereotypic meanings, and references (Morrone, Declercq, Novella, & Besche,

2010; Radvansky et al., 2010; Radvansky et al., 2009; Uekermann, Thoma, & Daum, 2008; von Hippel, Silver, & Lynch, 2000). It has been suggested that the difficulties in discourse comprehension are related to a decline in cognitive processes that have occurred with ageing (Brébion, 2003; Hannon & Daneman, 2009; Morrone et al., 2010; Radvansky et al., 2010; Radvansky et al., 2009; Uekermann et al., 2008; von Hippel et al., 2000). For instance, declines in cognitive abilities such as working memory capacity and executive processing have been noted (Finnigan, O'Connell, Cummins, Broughton, & Robertson, 2011).

Although there is general agreement that age-related cognitive decline and discourse comprehension difficulties are related (Brébion, 2003; Krawietz, Tamplin, & Radvansky, 2012; Radvansky et al., 2010; Radvansky et al., 2009), the nature of the relationship is unclear. For instance, one theory is that as working memory capacity declines, older individuals are unable to *store* information in order to process it (Brébion, 2003). In contrast, others have suggested that it is the processing of information that is impaired. For instance several researchers have proposed that inefficient inhibitory function in older people might account for difficulties in discourse comprehension processing (Morrone et al., 2010; Radvansky et al., 2010; Radvansky et al., 2009; Uekermann et al., 2008; von Hippel et al., 2000). That is, older adults may be less able to suppress unrelated contextual information which, in turn, makes inferring the correct meaning challenging when different interpretations are possible for a single linguistic form (Radvansky et al., 2010; Radvansky et al., 2009).

Anaphora resolution is one linguistic process involved in discourse comprehension that requires both the ability to store information as well as process it through the suppression of unrelated information. As an example, in the sentence “Frank loaned Jerry a pen but he wanted it back before long”, the information in the first clause needs to be maintained until

the unrelated information is determined by the anaphoric pronoun "he". The irrelevant information, which is the nonreferent name "Jerry" in this example, then needs to be suppressed so that "Frank" can be recognised as the single referent of the anaphoric pronoun and thus anaphora is resolved. Studies have supported the notion that anaphora resolution may be affected by age-related changes in cognitive abilities such as working memory capacity and inhibitory abilities[4] (Cohen, 1979; Light & Capps, 1986; Ulatowska et al., 1986); however, there has been minimal research on how the processing of potentially ambiguous anaphora is affected by age-related changes in different working memory functions. Hence, this research aimed to investigate younger and older participants' resolution of potentially ambiguous anaphora under low and high working memory load conditions, and to determine whether any age-related differences were related to individual differences in working memory capacity.

The purpose of this introductory chapter is to: 1) provide a review of the research on age-related changes in language comprehension, in particular discourse comprehension; 2) detail the concept of working memory and highlight seminal models and theories of working memory; 3) present a short review of the salient theories that describe the relationship between working memory and language comprehension; 4) describe the process of discourse comprehension and reference resolution in the framework of structure building; and 5) detail the research aims of the current thesis.

1.2 AGE-RELATED CHANGES IN LANGUAGE COMPREHENSION

[4] Suppression and inhibition has commonly been used interchangeably in the literature. However, for the purposes of this thesis, *suppression* is defined as the process of decreasing the accessibility of previously-relevant information which is already activated in working memory while *inhibition* is defined as the process of preventing the unrelated information from becoming activated. *Inhibitory abilities* is used as a general term to refer to different executive functions responsible for inhibiting a response or information by preventing it from becoming activated, or suppression of already-activated information. Further discussion of these concepts can be found on page 9.

Language comprehension plays an important role in social interactions throughout the lifespan. Language comprehension is reported in several studies to be negatively affected by ageing (Au et al., 1989; Brébion, 2003; Cohen, 1979; Federmeier & Kutas, 2005; C. L. Lee & Federmeier, 2012; Obler, Fein, Nicholas, & Albert, 1991; Radvansky et al., 2010; Radvansky et al., 2009; Uekermann et al., 2008; Ulatowska et al., 1986) resulting in older adults' loss of interest in social interactions (Burke & Shafto, 2008).

Not all language comprehension tasks are equally affected by age. For instance, older adults demonstrate few age-related changes on single-word comprehension tasks (Federmeier et al., 2003; Giaquinto et al., 2007). In contrast, studies of sentence comprehension have revealed age-related decline in processing speed (Norman, Kemper, & Kynette, 1992; Wingfield et al., 2003), in comprehending sentences under distracting conditions as in the presence of background noise (Wingfield et al., 2003), and in the comprehension of syntactically complex sentences (Norman et al., 1992; Wingfield et al., 2003). More recently, age-related changes have been revealed in the comprehension of linguistic segments larger than single sentences, i.e. *discourse* (Federmeier & Kutas, 2005; Federmeier et al., 2003).

1.2.1 Discourse Comprehension in Ageing

Discourse refers to the connected speech which is the main mode of daily communication. It includes more than just a sequence of individual words and sentences. Discourse is used in everyday conversation as well as in written materials. It involves integration of previous and new linguistic information with real-world knowledge and experience of the readers or listeners (Graesser, Millis, & Zwaan, 1997). Linguistic and real-world information are integrated at several levels of discourse processing to form a text's meaning. Age-related changes have been reported in different levels of discourse processing.

One of the most commonly accepted models of discourse processing is that proposed by van Dijk and Kintsch (1983). According to van Dijk and Kintsch (1983), discourse is processed as follows: surface level, text-based level, and situation model level. The surface level of discourse occurs when the meanings of the words and syntactic structures are decoded to form a surface representation of the discourse. The text-based level of discourse reflects the inter-connections and inferences that exist at text-level. For instance, pronoun-referencing can be an example of text-based processing if the pronoun's referent can be determined merely based on syntactic relations such as gender and number agreements (e.g., *John handed Sarah some tickets to a concert but he took the tickets back immediately*.). However, if the pronoun is potentially ambiguous and its syntactic features match more than one potential referents (e.g., *Bill handed John some tickets to a concert but he took the tickets back immediately.*), the pronoun is required to be further processed at situation model level (see Section 1.5.1 for more details about reference resolution). The situation model which is the third level of discourse processing is a mental image of what the discourse is about. The situation model includes an individual's previous experience and knowledge of the world in addition to the text meanings. The situation model is updated when new information is received (van Dijk & Kintsch, 1983).

Studies have examined the effects of ageing against the van Dijk and Kintsch (1983) model of discourse comprehension—the results have been equivocal. For instance, Radvansky and colleagues (1999, 2001) argued that discourse comprehension difficulties in older adults were results of surface-level impairments, such as remembering the words or syntactic relations of sentences. Likewise, these authors and others argued that situational levels of discourse processing involving knowledge integration and context use were preserved in ageing (Radvansky, 1999; Radvansky, Copeland, Berish, & Dijkstra, 2003; Radvansky & Dijkstra, 2007; Radvansky, Zwaan, Curiel, & Copeland, 2001).

In contrast, others have reported the opposite; noting changes in discourse comprehension at situation model levels of processing. Those changes include declines in the ability to follow the discourse characters and to use contextual information to make predictions and inferences (Federmeier & Kutas, 2005; Federmeier et al., 2003; Noh & Stine-Morrow, 2009). One area of discourse processing that has been more consistently reported to decline with ageing is situational-level discourse; such as the ability to infer implicit meanings in discourse (Borella et al., 2011; Cohen, 1979; Light & Capps, 1986; Ulatowska et al., 1986). These situational-level difficulties include age-related decline in comprehending metaphors, stereotypic meanings, and potentially ambiguous references in the discourse (Morrone et al., 2010; Radvansky et al., 2010; Radvansky et al., 2009; Uekermann et al., 2008; von Hippel et al., 2000).

In order to understand the nature of discourse comprehension decline with ageing, it is critical to not only understand levels of discourse processing but to understand the cognitive processes involved in comprehension at each level. It has been suggested that at least some of the mechanisms and processing underlying language comprehension involve general cognitive processes such as working memory, attention and inhibitory abilities (Gernsbacher, 1997b). Therefore, theories of cognitive ageing have attempted to explain the underlying mechanisms of age-related decline in language abilities (Brébion, 2003; Hasher & Zacks, 1988; Salthouse, 1996). The most salient of these theories are reviewed in the following section.

1.2.2 Theories of Cognitive Ageing and Language Decline

Cognitive function is commonly affected in ageing. However, since not all cognitive functions decline with ageing, the nature of age-related cognitive changes has remained controversial. Debate continues over whether the age-related changes in language

comprehension are purely linguistic (Au et al., 1989; Fodor, 1983) or due to a more general cognitive decline (see the review in Kensinger, 2009). The coexistence of linguistic deficits with other cognitive disorders in ageing has led some scholars to suggest that language decline is mediated by declines in more general cognitive functions (see the review by Goral, 2004).

The most commonly proposed causes underlying cognitive ageing include a decline in general processing speed (Salthouse, 1996), working memory capacity (Borella et al., 2011; Brébion, 2003; De Beni, Borella, & Carretti, 2007), and inefficient executive processing such as inhibitory decline (Hasher & Zacks, 1988). The underlying mechanisms of language processing difficulties in older adults have been commonly defined using these three theories. These theories mainly assumed that language processing required the use of limited cognitive resources and that exceeding this limitation resulted in language difficulties. The theories of general slowing, inhibitory decline, and working memory capacity decline are introduced in the following sections.

1.2.2.1 General Slowing Theory of Cognitive Ageing

The general slowing theory was developed by Salthouse (1996), based on the belief that older adults were slower than younger adults in both motor and cognitive functions. Salthouse (1992, 1996) suggested that older adults' poorer performance on cognitive tasks resulted from a slowed rate of processing. This slower processing speed resulted in less efficient use of limited cognitive resources. Based on this theory, older adults are proposed to devote a longer time to language processing—resulting in the loss of the earlier information. Moreover, older adults' performance is suggested to be poorer in tasks in which the rate of receiving the information can not be self-controlled. Listening comprehension is an example of such tasks, as listeners can not control the speakers' speech rate. In such case, spending

longer times on processing would require simultaneous processing of more information. That is because the new information is presented to the listener regardless of whether the processing of earlier information has been completed or not. Higher processing demand in such tasks may therefore exceed the limit of cognitive resources and result in comprehension difficulties. Unlike general slowing theory, some theories of cognitive ageing have focused on changes in single cognitive functions such as inhibition.

1.2.2.2 Inhibitory Decline Theory of Cognitive Ageing

The inhibitory decline theory of cognitive ageing suggests that older adults have difficulties inhibiting unrelated information while focusing on the goal-related information. For instance, this theory would predict older adults to be less accurate in performing a comprehension task in the presence of background noise, as they might be less able to inhibit the background noise and focus on comprehending the text, which is the goal of the task. Inhibitory decline theory was first proposed by Hasher and Zacks (1988) based on a previous finding that older adults kept more information in their mind while performing comprehension tasks as they were unable to inhibit the irrelevant information (Hamm & Hasher, 1992). Using a probe recognition task to study inference making during comprehension, Hamm and Hasher (1992) found that older adults kept incorrect inferences along with the correct ones while younger adults only maintained the correct inference.

The inhibitory decline theory has been supported by a number of behavioural as well as neuroimaging studies on ageing (C. L. Lee & Federmeier, 2012; Malmstrom & LaVoie, 2002; Noh & Stine-Morrow, 2009; Radvansky et al., 2010; Radvansky et al., 2009). However, controversies exist on the nature of the inhibitory abilities that decline with ageing. While inhibition was traditionally viewed as a single ability, more recently studies have proposed that different types of inhibitory abilities might exist. Lustig, Hasher and Zacks

(2007) distinguished between inhibitory functions that controlled which information was to be activated, suppressed the accessibility of already-activated irrelevant information, and inhibited a strong but inappropriate response (Lustig, Hasher, & Zacks, 2007). Similarly, Gernsbacher (1997) distinguished between inhibition and suppression suggesting that *inhibition* was the process of preventing the information from being activated in working memory while *suppression* was the process of decreasing the accessibility of previously activated information (Gernsbacher, 1997b). Friedman & Miyake (2004b) have also differentiated among the inhibotory processes applied to previously relevant information, and proponent responses or interfereing distractors (Friedman & Miyake, 2004b).

Following the belief that different inhibitory functions might exist, some studies have suggested that the application of the inhibitory deficit theory of ageing to age-related comprehension difficulties was restricted to certain types of inhibitory abilities. These inhibitory abilities were those responsible for the suppression of the information that used to be relevant to the goals of the comprehension task but became irrelevant with increasing information (Bell, Buchner, & Mund, 2008). Not all inhibitory functions contribute equally to language processing. Since different goal-related inhibitory functions are suggested to be involved in cognitive processing, measuring inhibitory abilities during a specific online task can best demonstrate the goal-related inhibitory processing involved in that task. In contrast to the inhibitory decline theory that has focused on inhibitory processes involved in language comprehension, some theories have highlighted the important roles of information storage and processes that contribute to the maintenance of relevant information in working memory.

1.2.2.3 Working Memory Capacity-based theories of Comprehension Decline in Ageing

Working memory is believed to be responsible for the temporary storage and processing of information. Capacity-based theories mainly assume that there is a limited

capacity for storing and maintaining information in working memory for later processing and that this capacity declines with ageing. In these theories, age-related capacity decline is considered to mediate age-related changes in other cognitive processes such as language comprehension. Studies have found correlation of comprehension abilities with measures of working memory storage as well as measures of processing abilities (Brébion, 2003). Brébion (2013) suggested that age-related comprehension decline was due to a decline in a general shared capacity for storage and processing.

Capacity-based theory has also been supported by the results from the studies demonstrating that task complexity affected comprehension accuracy and processing rate in older adults more than it did in younger adults (Stine-Morrow, Ryan, & Leonard, 2000; Wingfield et al., 2003). These studies have argued that since older adults had more limited capacity for storing and processing of information, storage was deficient under higher task's processing demands. In order to understand working memory capacity-based theories of language comprehension decline in ageing, it is important to look at models of working memory in detail. The following section provides definition and a review of the seminal models of working memory.

1.3 WORKING MEMORY

Working memory can be defined as a system responsible for temporary storing and processing of information. It is critical to language processing, particularly discourse processing (Baddeley, 2003; Daneman & Carpenter, 1980). Listeners and readers keep information activated in their working memory to ensure that it can be processed later to comprehend further text meanings (Gernsbacher & Kaschak, 2006). Therefore, working memory storage and processing are expected to be crucial in keeping the related information active for further integrations and inferences.

Decline in working memory abilities has been commonly considered a primary deficit in the theories of ageing. These theories have argued that age-related changes in cognitive abilities such as language processing are mediated by a decline in more general cognitive functions such as working memory storage or its executive functions (Brébion, 2003; Hasher & Zacks, 1988). Studies investigating the contributions of working memory abilities to language processing in both younger and older adults have been influenced by how they defined working memory. Several models of working memory have been proposed in the literature. The following section briefly introduces the most salient models suggested for working memory.

1.3.1 Models of Working Memory

Early models defined working memory as a short-term storage system (Atkinson & Shiffrin, 1968; Waugh & Norman, 1965). They first evolved from the bi-component models of memory. These models argued that memory was a system consisting of primary and secondary storages. Primary storage was a limited-capacity short-term system for storing the information while secondary storage was a long-term system. Primary storage was later referred to as short-term memory (STM). Information was believed to be transferred from STM into the secondary storage called long-term memory (LTM) (Waugh & Norman, 1965). These memory models were further expanded to form the multi-store models. Studies on the roles of working memory in language processing that were based on these models mainly investigated how language processing was affected by the ability to temporarily store linguistic information. Later models of working memory argued that working memory included a processing component – in addition to being a short-term store. One of the most salient models that included a processing component was the multicomponent model introduced by Baddeley and Hitch (1974).

The multicomponent model of working memory considered working memory as a system that both stored and processed information. According to Baddeley and Hitch (1974), working memory consisted of two specialized storage components and one processing component. Based on this model, the "phonological loop" and "visuo-spatial sketchpad" were two storage systems responsible for storing, and maintaining verbal-acoustic and visual information respectively. Maintenance of information was suggested to be achieved through rehearsal within the storage components as well as attentional and inhibitory processes performed by a central executive. The central executive was believed to be responsible for attentional control (Baddeley, 2003, 2012). Central executive functions in this model were those functions that played the most important role in focusing attention, dividing attention in multiple tasks, and switching between tasks (Baddeley, 2012). A fourth component, the "episodic buffer" was also later added by Baddeley (2000) to the multicomponent model of working memory. The episodic buffer was suggested to be the linking component between different components of working memory, and between working memory, LTM and perception. The episodic buffer was capable of holding different types of information (e.g. verbal, visual), which was crucial in linking the different components. This buffer was believed to be limited in the amount of information it could hold at a time (Baddeley, 2012).

Applied to the study of language processing, the multicomponent model suggested that visual and auditory language-related information were stored in working memory's separate storage components. Rehearsal processes within these components, as well as central executive functions of attentional control were responsible for preventing the information from being lost while inhibitory processes were responsible for dropping the irrelevant information such as similar articulations or meanings (Baddeley, 2010, 2012). Based on this model, age-related changes in language abilities could be due to separate deficits in executive processing and storage of linguistic information. Later models of working memory directed

their focus to the central executive functions of working memory. They mainly defined working memory as the processes important in keeping information active for further processing (Cowan, 1999; Engle, Tuholski, Laughlin, & Conway, 1999).

Engle et al. (1999) defined working memory as the activated subset of the stored information and attentional processes. They highlighted the roles of processes necessary to control attention and inhibit distracting information in working memory. Engle et al.'s theory focused on the inhibitory and attentional functions and considered these processes as the core executive functions of working memory. Working memory capacity in their theory was viewed as the ability to keep information active in conditions involving distractions and interferences. According to this model, working memory capacity was the amount of inhibitory and attentional processing that could be applied to the activated information (Engle et al., 1999).

Similarly, Cowan (1999) focused on the attentional functions of working memory. Cowan's (1999) model mainly differed from the multicomponent model in that Cowan did not consider working memory as a system consisting of specialized storage systems separated from long-term memory. Based on Cowan's embedded-processes model, working memory referred to a subset of the stored information in the long-term memory that was activated and in the scope of attention. According to this model, the amount of information that could be kept in the activated subset at each time was limited which resulted in a limited capacity for working memory (Cowan, 1999). Based on these models, studies have attributed differences in language processing abilities to different inhibitory and attentional abilities.

Baddeley (2012) argued that these later theories shared their main assumptions and mostly differed in their terminology and focus. In spite of the fact that slight differences exist in definitions provided for working memory capacity, limited capacity has been one area of

agreement among theories of working memory. Studies have demonstrated that working memory capacity differed across individuals. Moreover, it was suggested that working memory capacity declines with normal ageing resulting in language processing difficulties (Brébion, 2003; Norman et al., 1992). As an example, Noh and Stine-Morrow (2009) investigated the contribution of age-related working memory capacity decline to the ability to follow discourse characters. Their findings suggested that smaller working memory capacity in older adults resulted in deficient attentional control and thus inability to follow the characters in the discourse (Noh & Stine-Morrow, 2009).

Working memory capacity has been reported to contribute to individual differences (Daneman & Carpenter, 1980) in language abilities and to underlie the age-related changes (Brébion, 2003) in language processing. However, different focuses of the theories of working memory have resulted in a controversy over the working memory functions underpinning individual differences in language abilities. Attempts to explain these individual differences led to the development of the individual difference-based theories of working memory. These theories are reviewed in the following section.

1.3.2 Individual Difference-Based Theories of Working Memory

The main goal of individual difference-based models was to explain why people differed in the ability to maintain information in their working memory under complex task conditions. Separate-resources theories have assumed that separate resources existed for storage and processing each with a limited capacity (Milton, 2008). According to Towse, Hitch and Hutton (2000), storage limits were independent of processing load and vice versa. Based on these theories, individual differences in maintaining information were due to differences in storage capacity (Towse, Hitch, & Hutton, 2000). Shared resources theories, however, believed in a shared capacity for storage and processing. They focused on the

processes involved in maintaining information in working memory and in making an efficient use of the limited capacity (Daneman & Carpenter, 1980; Engle, Cantor, & Carullo, 1992).

These theories first evolved from the study of Daneman and Carpenter (1980) that found a high correlation between scores on a reading span test and performance in different language comprehension tasks. Daneman and Carpenter (1980) believed that working memory had a limited capacity shared between executive processing and storage. Working memory capacity measured with the reading span task was an index of this shared capacity. Daneman and Carpenter (1980) suggested that there was a trade-off between storage and processing and that the processing demands received priority. Therefore, under high processing load of language comprehension, storage would be deficient in people with lower working memory capacity. Older adults' language difficulties have also been explained using the trade-off theory (see Section 1.2.2.3 for more details).

Engle et al. (1999) have focused on efficiency in the use of the limited shared capacity. They suggested that the individual differences in cognitive abilities were due to differences in the amount of the inhibitory and attentional processes that could be applied to the information in working memory (Engle et al., 1999). Baddeley (2012) has argued that most of the individual difference-based theories of working memory mainly focused on the cognitive processes —e.g. inhibition, attention— which were functions of the central executive in Baddeley and Hitch's multiple component model, and also accounted for the contribution of storage components (Baddeley, 2012). The same explanation for poorer language performance in people with lower working memory capacity have also been applied to comprehension decline in ageing.

In summary, working memory capacity is suggested by several studies to account for the differences observed in cognitive abilities including language comprehension. Capacity-

based theories of cognitive ageing presented in section 1.2.2.3 have attributed age-related changes in language abilities to a decline in working memory capacity. However, a decline in specific working memory functions such as inhibition rather than of a general capacity could also account for age-related decline in language processing (see Section 1.2.2.2). Working memory storage and its executive functions are all demonstrated to play important roles in language processing (Carpenter, Miyake, & Just, 1995; Daneman & Carpenter, 1980; Gernsbacher & Kaschak, 2006; Newman, Malaia, Seo, & Cheng, 2013). It is, however, still unclear how differences in different working memory abilities contribute to differences in language comprehension abilities in younger and older adults. A number of studies have investigated the roles of different working memory functions in the underlying mechanisms involved in language processing. The following section reviews comprehension studies that have addressed the roles of working memory in language processing.

1.3.3 Working Memory and Language Comprehension

Most behavioural studies examining language comprehension and working memory have investigated correlations between measures of working memory and performance on language comprehension tasks. Different measures of working memory abilities have been used to study the relations between working memory and comprehension. Traditional measures of working memory that mainly examined the ability to store information in working memory did not correlate highly with performance on comprehension tasks (Carretti, Borella, Cornoldi, & De Beni, 2009). However, scores on more complex tasks involving both processing and storage (e.g. Reading Span Test designed for measuring working memory capacity) have revealed a correlation of working memory capacity and comprehension performance in both younger and older adults (Daneman & Carpenter, 1980; Noh & Stine-Morrow, 2009).

Using the Reading Span Task to measure working memory capacity, Daneman and Carpenter (1980) showed that comprehension was affected by referential distance (the number of sentences between the referent and the pronoun reference) in people with smaller working memory capacity. Moreover, those with smaller working memory capacity had lower comprehension scores compared to those with larger working memory capacity. Since the reading span task mainly measured storage and maintenance abilities (Was, Rawson, Bailey, & Dunlosky, 2011), these abilities were suggested to play important roles in language processing (Daneman & Carpenter, 1980). Following Daneman and Carpenter, Just and Carpenter (1992) suggested that language comprehension abilities depended on individual differences in working memory capacity. They proposed that information became activated during comprehension by processing written or spoken text and was stored in working memory if their activation level was higher than a threshold. Maintaining the information in working memory is important during language comprehension as the information might be updated, integrated to further text information or previous knowledge, and used to infer implicit information to form the comprehension output. However, there is a limit to the amount of activations in working memory. Therefore, the information in working memory might be lost due to either activation decay in time or displacement. When added information exceeds working memory capacity, previous information might be displaced by the newer information.

Based on Just and Carpenter (1992), working memory processes were responsible for preventing the loss of information through changing the levels of activations. Just and Carpenter (1992) argued that many of the processes involved in comprehension might occur simultaneously resulting in various partial products (activations). If the processing demand of a task were high, the amount of activations of the partial products would exceed activation limits of working memory. Therefore, some partial products might be lost in people with

lower working memory capacity. Just and Carpenter (1992) examined the effect of tasks' working memory demand on language comprehension. They increased working memory demand through increasing syntactic complexity (use of object relative vs. subject relative sentences, e.g., *the senator who the reporter attacked was warned by the policeman*.), increasing referential distance —i.e. the number of intervening clauses between a pronoun and its referent—, and adding ambiguity (e.g., *the senator attacked by the reporter admitted the error.*). Results from their study suggested that people with higher working memory capacity were faster and more accurate than people with lower working memory capacity in comprehending high demanding sentences.

Although working memory capacity as measured with the Reading Span Task has been found to positively correlate with comprehension abilities, the underlying association has remained controversial (Friedman & Miyake, 2004a). Proponents of the separate-resources theories of working memory capacity have proposed other reasons for correlation of comprehension with reading span scores, such as time-based forgetting (high processing requires more time which results in forgetting information) (Towse et al., 2000). Moreover, there is some evidence that individual differences in language comprehension are mainly due to different abilities in specific comprehension-related processes of working memory. Studies have demonstrated that individual differences in comprehension were associated with difficulty in working memory updating process, particularly in inefficient suppression of unrelated discourse information (Carretti, Cornoldi, De Beni, & Romanò, 2005; Gernsbacher, 1997b; Gernsbacher & Robertson, 1999; Gernsbacher, Varner, & Faust, 1990). The same controversy exists over the age-related changes in language comprehension. It is now most commonly accepted that older adults have more limited working memory capacity and that individual differences in working memory capacity contributes to language comprehension abilities. However, it is not yet clear whether age-related changes in language comprehension

are merely due to differences in working memory capacity or other age-related cognitive changes are involved.

As discussed in Section 1.2, age-related language difficulties have been mainly reported in discourse-level processing. Therefore a number of studies have been devoted to investigating the relationship between working memory abilities and discourse processing (Federmeier et al., 2003; Radvansky, 1999; Radvansky et al., 2003; Radvansky et al., 2001). Such studies have led to the development of cognitive models of discourse. These models have assumed that general cognitive abilities were involved in the comprehension of discourse (see Graesser et al., 1997 for a review of these models). A well-documented model, which has highlighted cognitive mechanisms involved in discourse processing, is the Structure Building Framework (SBF). This framework has explained the roles of cognitive processes in discourse comprehension and provided a theoretical framework in which to examine the individual differences in working memory and comprehension. This framework is briefly introduced in the following section with a focus on how the relationships between working memory abilities and discourse processing in ageing can be examined in this framework.

1.4 STRUCTURE BUILDING FRAMEWORK

Structure Building Framework (SBF) (Gernsbacher 1990) refers to a model proposed to explain the underlying mechanisms involved in discourse comprehension. It is based on the assumption that non-linguistic general cognitive processes were involved in the underlying mechanisms of language comprehension. Cognitive processes suggested in SBF to underlie language processing are mainly functions of working memory. Therefore, contributions of age-related working memory decline to discourse processing can be studied in this framework.

Based on SBF, the goal of language comprehension is to build a coherent mental representation or a mental structure of the discourse that is being processed. This mental representation corresponds to the situation model representation (van Dijk & Kintsch, 1983) of discourse. In the SBF, building the mental representation of comprehension included three sub-processes; laying a foundation, mapping relevant information, and shifting to a new substructure. Discourse concepts are represented in the mental structure through any of the three sub-processes. Therefore, discourse concepts could have different representations depending on the sub-process applied to them. The existence of different representations has been assumed by Gernsbacher to account for many comprehension phenomenon, such as *the advantage of first mention*. The advantage of first mention is that the first mentioned participants in a sentence, are more memorable and more highly accessible than other participants mentioned later in the same sentence as they form the foundation of the mental representation (Gernsbacher & Foertsch, 1999). Such comprehension phenomenon reflected outputs of the different stages of discourse processing in SBF. Investigating the effects of ageing on such phenomenon can thus provide insight into age-related changes in the process of discourse comprehension and different stages of building the mental representation of comprehension. Moreover, in SBF, the cognitive processes involved in different stages of language processing are mainly functions of working memory. Therefore, examining how differences in working memory abilities affect outputs of different sub-processes of language comprehension can shed lights on the contributions of working memory decline to language comprehension difficulties in ageing.

1.4.1 Contributions of Working Memory to Discourse Processing in SBF

The assumptions in SBF are consistent with theories of cognitive ageing as well as the individual difference-based theories of working memory and language comprehension. SBF

assumed that the cognitive processes underlying discourse comprehension were general cognitive processes. Therefore, the differences observed in language comprehension abilities among individuals might not always be language-specific, but instead might result from individual differences in cognitive abilities (Gernsbacher et al., 1990). SBF allows for the investigation of individual differences in the cognitive abilities that are involved in language comprehension.

Gernsbacher (1990) considered the building blocks of the mental representation to be memory cells or "memory nodes". The first activations of these memory nodes formed the foundation of the structure. This foundation has been demonstrated to be more accessible and more highly activated than other discourse information, as further information needed to be added onto it (Gernsbacher, 1990). According to the models of working memory and language processing (Baddeley, 2003; Baddeley & Hitch, 1974; Engle et al., 1999; Just & Carpenter, 1992), the activated information is stored in working memory until used for further processing and comprehension of meaning. It would thus be expected that working memory capacity affect the ability to build this foundation as it is an index of individuals' ability to store the information in working memory and to keep it activated in the presence of distractions. In SBF, the process of building the foundation has been demonstrated through the advantage of first mention (see Section 1.5.1.2. for detailed explanation and discussion). The contributions of individual differences in working memory to the process of laying the foundation is thus expected to be observed on the phenomenon of the advantage of first mention.

As further information is received, the mental representation is updated. If the new information was related to the previous information, it was mapped onto the previous mental representation as it reactivated the similar nodes. However, if the incoming information was

irrelevant or less relevant to the previous structure, it activated different nodes and the comprehenders shifted to build a new substructure (Gernsbacher, 1990; also see Giora, 1996 for a comprehensive review of SBF). Since working memory capacity is limited, the process of update and the addition of new information require older information to be suppressed from working memory. Besides, older information that might be needed for further processing might decay in time (Just & Carpenter, 1992). Therefore, the accessibility of the relevant information needs to be enhanced while distractions need to be inhibited.

Executive functions of working memory are responsible for controlling which information is to be attended, inhibited from being activated or ignored by being suppressed (Baddeley & Hitch, 1974; Engle et al., 1999). Therefore, the cognitive processes of decreasing and increasing the accessibility of information in SBF are mainly associated with working memory central executive which is responsible for attentional and inhibitory functions (Carroll, 2008). Since suppression and enhancement, which are processed by working memory central executive, play a main role in discourse processing, it is expected that individual differences in these executive functions affect comprehension. Gernsbacher focused on suppression skills as the main source of individual differences in comprehension abilities. The author argued that individual differences in discourse comprehension might reflect differences in suppression abilities (Gernsbacher, 1997b).

In a number of studies, Gernsbacher and colleagues have demonstrated that people with low and high language comprehension scores were equally able to activate potential meanings in the discourse. However, people with lower comprehension accuracies were less able to suppress the unrelated discourse information compared to people with higher comprehension accuracies (Gernsbacher, 1997b; Gernsbacher & Faust, 1991; Gernsbacher, Keysar, Robertson, & Werner, 2001; Gernsbacher & Robertson, 1999). Activation levels in

Gernsbacher and colleagues' studies were measured using a probe-response paradigm. In their paradigm, participants were required to complete two tasks. Firstly, they were presented with written or spoken discourse to comprehend. Secondly, while comprehending the discourse, participants were presented with a probe that was a word from the presented discourse or a potential inference-based meaning of the discourse. How long participants took to recognize the probe represented how activated the probe was in their mind. The more activated a probe was, the more accessible it was in comprehenders' mental representation of comprehension. Gernsbacher assumed that the changes that occurred in the accessibility of discourse information reflected the online process of comprehension (Gernsbacher, 1989, 1990, 1991).

In summary, SBF suggested that discourse processing involved laying foundations for mental representation of discourse by storing information and maintaining them in working memory, adding more information, and updating this foundation by suppressing no-longer relevant information and enhancing the relevant information. Working memory is responsible for attentional control to maintain information active or suppress the accessibility of the information. It is thus expected that differences in working memory abilities affect discourse processing. There is evidence that working memory abilities are negatively affected by ageing resulting in a decline in comprehension abilities (Finnigan, O'Connell, Cummins, Broughton, & Robertson, 2011; Radvansky et al., 2009). Investigating these processes during discourse comprehension in older and younger adults can thus provide insight into the underlying mechanisms involved in age-related changes in discourse comprehension.

1.4.2 Cognitive Ageing in Structure Building Framework

Considering the SBF's account of discourse processing, both capacity-based and inhibitory decline theories of ageing can explain the underlying mechanisms for age-related

changes in discourse comprehension. Based on a capacity-based theory of cognitive ageing (Brébion, 2003), it would be expected that the age-related discourse processing decline be mainly due to a decline in general capacity for storing and maintaining the activated information in working memory. Inhibitory decline theory of ageing (Hasher & Zacks, 1988) would argue that discourse comprehension difficulties in ageing are primarily caused by a decline in inhibitory abilities. Contributions of working memory capacity and inhibitory abilities to age-related changes in discourse comprehension can be investigated in SBF through examining the effect of age on the advantage of first mention and the suppression process. Previous studies have managed to document these processes in a number of discourse processing (Carreiras, Gernsbacher, & Villa, 1995; Gernsbacher, 1989, 1991, 1997a, 1997b; Gernsbacher & Faust, 1991; Gernsbacher & Hargreaves, 1988; Gernsbacher & Jescheniak, 1995; Gernsbacher et al., 2001; Gernsbacher & Robertson, 1999; Gernsbacher et al., 1990).

One aspect of discourse processing during which both suppression process and the advantage of first mention have been demonstrated is anaphora resolution. Anaphora resolution is the process of finding the referent of an anaphora. Anaphora is a commonly-used linguistic device in English which contributes to the continuity of discourse by refereeing back to previously mentioned discourse entities (Graesser et al., 1997). Successful anaphora resolution is important in discourse comprehension as anaphora plays an important role in keeping discourse connected and coherent. Anaphora resolution ability has been commonly reported to decline with ageing resulting in communication difficulties (Borella et al., 2011; Cohen, 1979; Light & Capps, 1986; Ulatowska et al., 1986). Storing and maintaining information in working memory as well as working memory executive processing have been suggested to play important roles in anaphora resolution (Gernsbacher, 1989). Therefore, study of anaphora resolution provides good source for obtaining

information about the age-related changes in discourse comprehension and contributions of working memory capacity and inhibitory decline to differences in comprehension abilities. However, a limited number of studies have investigated the effects of ageing on the process of anaphora resolution.

1.5 ANAPHORA RESOLUTION AND WORKING MEMORY IN AGEING

Of the few studies on anaphora resolution in ageing and its correlation with working memory decline, most have been offline studies using working memory capacity and information recall tasks, and the results were mixed. Some scholars have considered a deficit in working memory storage to be the main reason for comprehension decline (Light & Capps, 1986). Light and Capps (1986) studied older adults' ability to find the referents for potentially ambiguous anaphoric pronouns under different working memory loads. No difference was found in the abilities of younger and older adults in comprehending anaphoric pronouns under lower memory load (Example 1 below). Under higher working memory load (Example 2 below), however, older adults performed worse than younger adults in resolving anaphora. In this study, working memory load was increased by increasing the referential distance. The ageing effect was noticeable when two intervening sentences were added between anaphora and its referent. Light and Capps (1986) expanded the experiment by asking the participants to recall the sentences at the end of each trial after finding the antecedents. They argued that older adults were as capable as younger adults in integrating the contextual and pragmatic information to resolve anaphora. However, they had difficulties remembering this information. When contextual information was forgotten, it could not be integrated to the further contextual information which could have resulted in a deficient anaphora resolution (Light & Capps, 1986).

Example 1: Henry spoke at a meeting while John drove to the beach. *He* brought along a surfboard.

Example 2: Henry spoke at a meeting while John drove to the beach. It was a nice day and there was the sound of activity in the streets. *He* brought along a surfboard.

A few studies have recently used an online approach to study the process of anaphora resolution in older adults. Using a gender stereotype mismatch paradigm, and measuring reading times, Radvansky, Lynchard, and Hippel (2009) investigated age-related differences in anaphora resolution. Experimental sentences used in this study introduced a character with his/her profession (e.g. babysitter) which was referred to later in the sentence using an anaphoric name reference. The professions used in the sentences had a gender stereotype (e.g. babysitter = female). Anaphoric names were either consistent or inconsistent with the profession stereotypic gender (e.g., babysitter was a young girl/boy who always took the job seriously and got along with Paul very well). Reading times of the anaphoric references were used to measure accessibility and expectancy of the potential inferences. Previous research had shown that young adults had longer eye-fixation times on references with stereotype-inconsistent gender when the gender had not been explicitly identified earlier (Duffy & Keir, 2004). This study had demonstrated that younger adults were able to resolve anaphora by suppressing the wrong inference. Radvansky et al.'s (2009) study on older adults demonstrated that older adults were more prone to the influence of stereotypic information and were less likely to suppress the incorrect stereotypes. Findings from this study were in line with the inhibitory decline theory of cognitive ageing suggesting that older adults were unable to suppress the accessibility of the irrelevant information in discourse.

Using eye-tracking in a series of expectancy violation experiments, Shake and Stine-Morrow (2011) examined age-related changes in anaphora resolution. Prior disambiguation

and presence of facilitating contextual information were manipulated factors in this study. The authors found that older adults relied more on contextual information to find the referent and had more regressive eye-movements to reprocess the discourse for finding the referent. They attributed the results to a greater allocation of cognitive resources to discourse-level information in older adults (Shake & Stine-Morrow, 2011). Although measuring reading times have been effective in examining age-related difference in processing anaphoric references, it could not demonstrate individual differences and age-related changes in specific cognitive processes involved in anaphora resolution. To investigate the contributions of different working memory abilities to age-related decline in the process of anaphora resolution, it is important to have an understanding of the cognitive mechanisms underlying anaphora resolution.

Structure Building Framework has been suggested to be able to account for cognitive processing underlying anaphora resolution (Gernsbacher, 1990). SBF has been supported by a large number of experiments including comprehension of different linguistic structures (Gernsbacher, 1989, 1990, 1997a, 1997b; Gernsbacher & Hargreaves, 1988; Gernsbacher et al., 2001; Gernsbacher & Robertson, 1999; Gernsbacher, Tallent, & Bolliger, 1999; Gernsbacher et al., 1990). As discussed in Section 1.4.1, in SBF, working memory capacity and executive functions play crucial roles in discourse processing. Therefore, Gernsbacher's (1989, 1990) model of anaphora resolution in SBF provides a good framework in which to study the contributions of age-related working memory decline to the changes in language comprehension.

1.5.1 Anaphora Resolution

Anaphora is the most common type of discourse reference in English. Anaphoric references are linguistic devices that play an important role in keeping discourse connected

by referring to discourse entities that have been previously mentioned in the discourse (Graesser et al., 1997). For example, in the sentence *John bought a hamburger for Sarah because she was hungry*, *she* is an anaphoric pronoun referring back to *Sarah*. *Sarah* is the *referent* of the anaphora. The process of determining whom/what is being referred to through anaphora is called *anaphora resolution*. Anaphora resolution is an important process in understanding discourse relations and meanings.

In some sentences, such as the above example, grammatical and semantic features (e.g. gender agreement, number agreement) of the anaphora match only one of the entities mentioned previously in the sentence. In some cases, however, they match more than one entity. In the sentence *Bill handed John some tickets to a concert but he took the tickets back immediately*, semantic and grammatical features of *he* match both proper names in the first clause; *Bill* and *John*. In such sentences, the anaphora can only be comprehended if other discourse information and the reader's previous knowledge of the word are integrated into the lexical meanings and features of the text. Thus, anaphora resolution usually requires more than just surface and text-based processing of the discourse. High-level processing should result in selecting a single referent among discourse entities that can potentially be the referents of the anaphora.

Each entity mentioned in a discourse, including the referents, has a degree of accessibility in the reader's or listener's mind. The degree of accessibility of each entity is measured relative to the accessibility of the other discourse entities, or to the accessibility of the previously stored information in the long-term memory. Anaphora is successfully resolved if its referent is sufficiently more accessible than other entities in working memory (Greene, McKoon, & Ratcliff, 1992; Streb, Hennighausen, & Rösler, 2004). According to Gernsbacher (1990), previously stored information in memory is activated by incoming

linguistic stimuli. Applied to anaphora resolution, when an individual hears an anaphoric pronoun (e.g. she) within a sentence, this pronoun activates the potential referents of the anaphora. A process of competition occurs until the target concept achieves greater activation than its competitor concepts—the end result of which is successful comprehension (Gernsbacher, 1990).

Two main cognitive processes are suggested by Gernsbacher (1989, 1990) to facilitate anaphora resolution—suppression and enhancement. Enhancement improves the accessibility of the referent by increasing its levels of activation, while suppression improves the accessibility by decreasing the activation of other irrelevant concepts in the discourse—nonreferents. Based on the SBF, the difference between the level of accessibility of referents and nonreferents is positively affected by the degree of explicitness of the anaphora (Gernsbacher, 1989). In a series of experiments, Gernsbacher (1989) showed that suppression and enhancement were triggered by anaphoric references in the discourse. The author investigated the concepts' levels of accessibility through measuring response times to referent and nonreferent names in a probe recognition task presented to participants while they were reading sentences containing potentially ambiguous anaphoric pronouns (e.g., *Bill handed John some tickets to a concert but he took the tickets back immediately.*).

1.5.1.1 Gernsbacher's (1989) Research into Anaphora Resolution

Gernsbacher (1989) investigated the roles of suppression and enhancement in the comprehension of anaphoric references in six experiments. It was hypothesized that reference resolution was facilitated by enhancing the accessibility of the referents (REF) and suppressing the accessibility of the nonreferent (NREF) concepts in the discourse. In each experiment, participants were asked to read a series of sentences for comprehension. Each sentence contained two main clauses; the first main clause introduced two characters using

proper names; in the second main clause, one of these names was referred to using either a noun or a pronoun anaphora (Example 1 below).

Example 1: Bill handed John some tickets to a concert but *Bill / he* (anaphora) took the tickets back immediately.

While reading the sentences, participants were presented with a probe recognition task. Three variables were manipulated in Gernsbacher's experiments. The first variable was probe type (REF/NREF). In each trial, the probe was either the referent of the anaphora or the nonreferent name in the first clause, or another name not present in the sentence. Readers were to decide whether they had seen the probe name in the sentence. Readers' response times (RTs) to the probe recognition task were measured to evaluate the degree of the accessibility of the referent and the nonreferent names. The longer it took the readers to recognize a probe, the less accessible it was in their mind. The second variable was the explicitness of anaphora (pronoun/noun anaphora). The third variable in Gernsbacher's experiments was the testing point for the probe recognition task (immediately prior to the conjunction / immediately before the anaphora / immediately after the anaphora / at the end of the sentence). Testing points changed across the experiments and in each experiment, two of the above testing points were used to measure changes in the accessibility of the referent and nonreferent names. First and second testing points in Gernsbacher's experiments 1-3 are presented in Table 1.1.

Table 1.1 : *Probe recognition testing points Gernsbacher's (1989) Experiments 1-3*

	First Main Clause	**TP**	**Conj**	**TP**	**NA/PA**	**TP**	**Second Main Clause**	**TP**
	Bill handed John some tickets to a concert		but		Bill / he		took the tickets back immediately.	
EXP1				■		■		
EXP2		■				■		
EXP3						■		■

Note: TP = Testing Point, Conj = Conjunction, EXP = Experiment, NA = Noun Anaphora, PA = Pronoun Anaphora, 1 = First Testing Point, 2 = Second Testing Point

In the first experiment, accessibility was measured immediately before and immediately after the anaphora. Results demonstrated that readers were faster at recognizing the referent probes compared to the nonreferent probes after noun anaphora. Reading a pronoun anaphora, however, did not change the accessibility of referent or nonreferent names. It was concluded that noun anaphora immediately triggered suppression and enhancement. These results were replicated in a second experiment in which the first testing point was moved back to the end of the first clause, just before the conjunction. In a third experiment, accessibility was measured right after the anaphora and at the end of the sentence. The results demonstrated that pronouns also suppressed the accessibility of nonreferents. Since suppression process was only observed at the end of the sentence, it was concluded that pronoun anaphora triggered suppression more slowly and less powerfully than explicit noun anaphora.

In a fourth experiment, Gernsbacher tested the possibility that slower suppression after anaphoric pronouns, compared to after a noun anaphora, was due to presence of semantic information after the anaphora and near the end of the sentence. In other words, this

experiment examined whether suppression was triggered by the information available in the anaphora or the semantic information in the sentence. For this experiment, new stimuli were used in which the semantic information was provided before the occurrence of anaphoric pronoun (Example 2 below). There were two sentences in each trial. The first sentence introduced two characters, one of which was referred to in the second sentence. The second sentence included a dependant clause providing the information required for anaphora resolution, followed by second clause containing the anaphoric reference.

Example 2: Bill lost a tennis match to John. Accepting the defeat, *he* walked quickly toward the showers.

Accessibility of the referent and nonreferent names were measured at the same points as in experiment 3; i.e. immediately after the anaphora and at the end of the sentence. Since the results replicated those from the third experiment it was suggested that suppression was not mainly triggered by the semantic information. The information from pronouns had to be combined with the semantic information in the second clause to allow anaphora resolution. With a fifth experiment (Example 3 below), Gernsbacher demonstrated that pronouns still triggered suppression more slowly even if they only matched the gender of one of the characters. In that case, however suppression was more powerful compared to when the pronoun matched the gender of both characters. In her last experiment, Gernsbacher showed that newly introduced participants also suppressed the accessibility of the other characters in the sentence (Example 4 below).

Example 3: Tim predicted that Tam would lose the track race, but *she* came in first very easily.

Example 4: Ann predicted that Pam would lose the track race, but *Sue* came in first very easily.

Gernsbacher's (1989) experiments demonstrated that both noun and pronoun anaphora triggered suppression which played the most important role in anaphora resolution. Providing a way to measure suppression abilities during online discourse processing, Gernsbacher's paradigm could be used to examine the inhibitory theory of cognitive ageing. This theory might predict an inability to suppress the irrelevant discourse information in older adults. Gernsbacher's (1989) study also provided a support for the advantage of first mention. Across all experiments, the first mentioned name was significantly more accessible than the second mentioned name (this advantage is explained in the following section). As mentioned earlier, this advantage is suggested by SBF to represent the process of laying the foundation for mental representation of discourse. Laying the foundation depends on the ability to store and maintain information in working memory so that further information can be mapped on to it (Gernsbacher, 1990). Working memory capacity as described by Daneman and Carpenter (1980) and Engle et al. (1999) is expected to influence the ability to lay foundations for comprehension, as it reflects the ability to maintain the information activated through attentional and inhibitory controls. Therefore, Gernsbacher's paradigm can be used to study how structure building process during anaphora resolution is affected by working memory capacity.

1.5.1.2 The Advantage of First Mention

Based on SBF, the order in which concepts are mentioned in a discourse affects their mental representation. It is believed that the first mentioned concepts in a sentence are the most accessible ones in the comprehenders' minds (Gernsbacher, 1990; Gernsbacher & Hargreaves, 1988). Gernsbacher and Hargreaves (1988) suggested that this advantage existed because the first concepts in the discourse formed the foundation of the mental representation, onto which the new information was to be mapped. Therefore, first concepts

were more accessible than others in a sentence (Gernsbacher & Hargreaves, 1988). Since this phenomenon is believed to demonstrate the process of laying the foundation during structure building, its investigation can shed lights on individual differences and age-related changes in language processing.

Gernsbacher and Hargreaves (1988) tested the advantage of first mention with a probe recognition task in two-character sentences. The authors also examined the possibility that the effect might be due to linguistic or semantic factors instead of the order of mention. The factors tested included, the semantic role of the character (agent or patient), syntactic role (subject or object), and whether or not the character was the sentence-beginning word. Their results showed no evidence for the mentioned effects. Although the same results have been observed in languages other than English (Carreiras et al., 1995; Kim, Lee, & Gernsbacher, 2004), there have been some disagreements regarding the proposed Advantage of the First Mention.

First, it has been argued that the advantage observed in Gernsbacher's experiments was due to task-specific strategies used in performing the probe recognition task rather than the processing involved in comprehension (Gordon, Hendrick, & Foster, 2000). However, further studies using other methods such as eye-tracking also provided evidence for the advantage of first mention during language comprehension (Järvikivi, van Gompel, Hyönä, & Bertram, 2005). Järvikivi and colleagues (2005) measured eye fixation times on the pictures of the potential referents of ambiguous pronouns. They found advantage for the first mentioned name compared to the second mentioned name, and for the subject of the sentence compared to the object of the sentence. These advantages were demonstrated by longer eye fixation times in this study (Järvikivi et al., 2005).

Second, it has been argued that the advantage of first mention was in contrast with the theory of the advantage of clause recency (Caplan, 1972). Based on the advantage of clause recency, words from the most recent clause were the most accessible words in mental representation. Gernsbacher, Hargreaves, and Beeman (1989) suggested that these two theories were not contradictory. They demonstrated that in complex sentences in which characters occurred in different clauses, the advantage of first mention was sensitive to the timeframe in which the accessibility was measured. The findings from Gernsbacher et al.'s (1989) study showed that in two-clause sentence containing a character in each clause (e.g. *Tina gathered the kindling and Lisa set up the tent*), the second character was more accessible immediately after offset of the second clause. This was explained by the advantage of clause recency (Caplan, 1972). However, when accessibility was measured with a delay, the first character became more accessible (Gernsbacher, Hargreaves, & Beeman, 1989). Gernsbacher et al. (1989) suggested that the clause recency advantage was short-lived and due to shifting to build a new substructure for the clause.

In summary, the advantage of first mention as well as the role of suppression in language processing have been well-studied and documented in English as well as in other languages (Gernsbacher, 1989, 1997a, 1997b, 1997c; Gernsbacher & Hargreaves, 1988; Gernsbacher et al., 1989; Gernsbacher et al., 2001; Gernsbacher & Robertson, 1999; Kim et al., 2004). While the observed advantage of the first mention reflected the process of laying the foundation, the observed suppression of the nonreferents demonstrated further processing of discourse and updating the mental representation. Investigating the age-related differences in these abilities can enhance our knowledge about the underlying mechanisms involved in language comprehension decline in ageing. However, a main controversy exists over the process of anaphora resolution that needs to be resolved if age-related changes in anaphora

resolution are to be studied. This controversy is over the time-course during which suppression of nonreferents occurs.

1.5.1.3 When is Anaphora Resolved?

As discussed earlier, Gernsbacher and colleagues have shown the resolution of pronoun anaphora through providing evidence that the accessibility of nonreferents was suppressed by the time readers finished reading the anaphoric sentences. Some studies, however, failed to find the suppression effect when accessibility of referents and nonreferents was measured at the end of the sentence (Greene et al., 1992; Love & McKoon, 2011). Thus, it had remained controversial when, in the process of anaphora resolution, the referent becomes more accessible relative to the other discourse concepts. This issue has raised a controversy over whether or not reference resolution is automatic and completed regardless of task demands.

It has been argued that most studies on pronoun resolution have taken an automatic memory-based approach to anaphora resolution. That is, these studies have implicitly assumed that anaphora was resolved automatically by selecting a single referent as soon as a reader encountered a pronoun (Greene et al., 1992). Based on these approaches, anaphora resolution was completed regardless of the task demands and the readers' strategies. In contrast, Greene, McKoon and Ratcliff (1992) believed that anaphora resolution was not always automatic, and could be strategic. They suggested that only the surface and text-based levels of anaphora processing were automatic while the rest depended on the task demands. McKoon and Ratcliff (1992) argued that semantic and grammatical features provided by an anaphora were matched automatically. If one discourse entity matched the pronoun features better than the others, it was automatically identified as the anaphora's referent. However, if more than one entity matched the pronoun features, the referent identification would not be

automatic. In such a case, resolution could be delayed under certain situations and depending on the task demands.

Love and McKoon (2011) suggested that referential distance contributed to the time-course for anaphora resolution. They showed that anaphora resolution was completed later in short passages than in longer discourse. Delayed resolution of anaphora in shorter texts in this study was attributed to partial engagement of readers in comprehension of short texts (Love & McKoon, 2011). Delayed anaphora resolution in shorter passages could be explained by the task's working memory demands. Since the amount of information in short passages was less than that of a longer passage, working memory demands were lower. Under the lower working memory demands, readers' working memory capacities were sufficient to keep all the information activated for further processing. Therefore, it is possible that since earlier suppression was not necessary, participants delayed the process until they were required by the comprehension question to resolve the anaphora.

However, under higher working memory demands, when much more information was presented to the readers, their working memory capacity might have not been sufficient to keep all the older information while simultaneously receiving and processing the new information. Therefore, older information could have been replaced with the newer information. To prevent the required information from being lost from working memory, readers might have had to suppress the irrelevant information so that the relevant information can be maintained. Therefore, it could be argued that under higher working memory load anaphora needed to be resolved earlier so that the irrelevant information could be suppressed. This is consistent with the finding from the study by Just and Carpenter (1992) that found an effect of working memory capacity on resolving ambiguity. This study showed that people with lower working memory capacity resolved ambiguity faster than those with higher

working memory capacity. They attributed this finding to the insufficient capacity for simultaneously maintaining the different interpretations of an ambiguous sentence and processing the remaining of the sentence. Therefore, while low capacity readers had to suppress the incorrect interpretation earlier to allow more capacity for further processing of the sentence, high capacity readers maintained all interpretations and delayed the disambiguation until it was required (Just & Carpenter, 1992).

If age-related decline in discourse comprehension is due to a decline in working memory capacity, it might be expected that ageing would affect the time-course for anaphora resolution. Older adults have more limited working memory capacity compared to younger adults. Based on the capacity-based theory of cognitive ageing (Brébion, 2003), more capacity would be devoted to processing than to storage in older adults which would result in deficient storage and loss or displacement of older information. In such case, it would be expected that older adults tend to suppress more information from working memory to allow more capacity for processing. Therefore, anaphora would be resolved earlier resulting in earlier suppression of the nonreferents, particularly under high storage load. Moreover, capacity-based theory would predict that increasing task's storage demands affect older adults' comprehension accuracy more than younger adults. In contrast, the inhibitory theory would consider the main reason underlying older adults' comprehension decline to be the inability to suppress the nonreferent. Therefore, regardless of tasks' storage demands, older adults would not suppress the nonreferent.

1.6 SUMMARY AND FUTURE DIRECTIONS

While discourse comprehension abilities are generally believed to decline with ageing (Federmeier & Kutas, 2005), the mechanisms underpinning these changes remain controversial. Only a limited number of studies have investigated the changes in the process

of discourse comprehension, and the results have been equivocal. A possible reason for the different findings in this field can be the different approaches taken to the study of comprehension abilities in ageing.

It is most commonly suggested that an age-related decline in more general cognitive abilities might mediate changes in language comprehension (Brébion, 2003; Hasher & Zacks, 1988; Salthouse, 1996), especially in the comprehension of ambiguous discourse concepts (Radvansky et al., 2010; Radvansky et al., 2009). Working memory capacity decline has been reported in several studies to result in discourse comprehension difficulties in older adults (Brébion, 2003; Norman et al., 1992). Although it is commonly accepted that ageing affects working memory capacity and that working memory capacity affects comprehension abilities, it is not yet clear whether age-related changes in language comprehension could only be attributed to working memory capacity decline. Changes in specific cognitive processes important for discourse comprehension (e.g. inhibitory processes) have been suggested to be the main reason underlying comprehension difficulties in ageing.

Previous studies on older adults' comprehension abilities have mainly used offline measures of discourse comprehension. While these offline studies were useful in examining the effects of age and working memory capacity on language comprehension accuracy, they were unable to differentiate between the changes caused by ageing and those due to working memory capacity. Online studies are required to investigate how the process of language comprehension is affected by these factors and whether their effects are independent of each other. However, the number of studies investigating the contribution of working memory capacity and executive processes to the online processing of discourse in older adults is limited.

Anaphora resolution, as a common discourse process which has been reported to be affected by ageing, has been the focus of a number of studies on the effects of working memory decline on language comprehension in older adults (Light & Capps, 1986; Radvansky et al., 2009). Although studies have suggested important roles for working memory functions in anaphora resolution, the association of age-related decline in working memory capacity and executive functions— e.g. inhibitory functions — with comprehension abilities have remained unclear. The existing literature on the age-related changes in underlying cognitive processes involved in anaphora resolution is limited. Moreover, there is still controversy over the time-course for resolving anaphora.

SBF (1990) provides a good theoretical framework in which to study the cognitive mechanisms involved in anaphora resolution. In particular, in SBF, effects of working memory and ageing can be studied on two important processes involved in anaphora resolution. First, the contribution of age and working memory capacity to the early process of laying the foundation for mental representation of comprehension can be examined through investigating any changes in the advantage of first mention. Secondly, the contribution of these factors to the process of updating the mental representation can be studied through examining any differences in the suppression of the irrelevant information.

It was suggested that while working memory capacity contributes to the ability to lay foundation for comprehension, executive process of suppression plays a crucial role in information update. As described earlier, Gernsbacher's (1989) paradigm for the study of anaphora resolution allows for the study of the processes involved in anaphora resolution. Adapting this paradigm for the study of anaphora resolution in older adults can thus shed lights on the age-related changes in discourse processing. Therefore, this research seeks to investigate the process of anaphora resolution in younger and older adults. It focuses on the

effects of working memory capacity and ageing on the advantage of first mention and the suppression of the irrelevant discourse information during anaphora resolution.

1.7 AIMS OF THE PRESENT THESIS

The overall aim of the current thesis was to investigate the underlying processes involved in anaphora resolution in younger and older adults. It examined whether anaphora resolution was affected by individuals' age and working memory capacity. The underlying processes involved in anaphora resolution were studied by an attempted replication and extension of Gernsbacher's (1989) study. Gernsbacher's paradigm also allowed for the investigation of comprehension-related suppression abilities in younger and older adults. The thesis is divided into three experimental chapters, with specific aims detailed below.

1.7.1 Chapters 3 and 4: effects of age and working memory capacity on anaphora resolution

Chapters 3 and 4 seek to study changes in the accuracy as well as the process of anaphora resolution in ageing. In particular, they investigate whether an age-related decline in working memory capacity contributes to changes in anaphora resolution in older adults. The questions addressed in Chapter 3 were: 1) Do age and working memory capacity affect anaphora resolution in a comprehension task?; 2) Do age and working memory affect advantage of first mention in a probe recognition task (as measured by two interactions among age, probe-type, and referent position, and working memory capacity, probe-type, and referent position)?; and 3) Does age affect suppression of irrelevant information in an anaphora resolution task (as measured by the interaction between probe type and probe testing point)?. Chapter 4 examined the same questions under higher working memory load. It also aimed to determine whether older adults' comprehension performance might be more negatively affected by increased working memory storage load of the task.

1.7.2 Chapter 5: a follow-up investigation into the process of anaphora resolution

Chapter 5 follow-ups on the findings from Chapters 3 and 4, and further examines the underlying processes involved in anaphora resolution and the factors that influence the time-course for suppression of the nonreferents. The specific questions of Chapter 5 are: 1) How is comprehension accuracy affected by working memory storage load and the syntactic structure of the sentences?; 2) Do younger adults suppress the accessibility of the nonreferents by the time they reach the end of the sentence, in simpler sentences with increased storage load and late disambiguation?; and, 3) Do younger adults suppress the accessibility of nonreferents by the time they reach the end of the sentence, in more syntactically complex sentences with low storage load and prior disambiguation?

2. CHAPTER TWO

Methodology

2.1 METHOD OVERVIEW

The current thesis is divided into three experimental chapters; Chapter 3: age-related changes in anaphora resolution, Chapter 4: age-related changes in anaphora resolution under higher working memory storage load, and Chapter 5: a follow-up investigation into the mechanisms that underlie anaphora resolution. Chapters 3 and 4 each includes a comprehension experiment and Chapter 5 includes four comprehension experiments. The experimental procedures employed were constant across the three chapters. However, across the experiments, a number of methodological and stimuli variables were manipulated. chapter-specific and experiment-specific information, including experimental procedure, participants and stimulus, are detailed in the relevant chapters.

2.2 PARTICIPANTS

One hundred individuals, aged 18 to 40 years, participated in this research programme: Chapters 3 and 4 ($n = 60$), Chapter 5 ($n = 40$). Participants in Chapters 3 and 4 included 30 younger (19-35 years old) and 30 older (66-87 years old) adults. Participants in Chapter 5 included 40 younger adults (18-38 years old). Mean age, standard deviation and gender information of the participants are presented in the relevant chapters. Participants were all right-handed native speakers of New Zealand English (NZE). They reported no history of neurological disease, dementia, cognitive impairment, cardiovascular disease, uncontrolled hypertension, learning disability, attention deficit disorder, or speech disorder.

The majority of the younger participants were recruited from the students at the University of Canterbury, Christchurch, New Zealand. Older participants were also recruited from the author's friends, local clubs and community organizations. Participants were assigned to research experiments in either Chapters 3 and 4 or Chapter 5 depending on time of recruitment. Individuals were recruited in October to December 2012 for the research

reported in Chapters 3 and 4 and in September to November 2013 for the study reported in Chapter 5.

Studies with similar paradigms measuring response times to probe recognition tasks in younger and older adults have found significant results with smaller numbers of participants (e.g. 24 participants in each group in Malmstrom & LaVoie, 2002). However, to increase the statistical power in the first experiment, we doubled the number of trials in each condition compared to Gernsbacher's (1989) study.

2.3 PROCEDURE

Testing was undertaken at the Department of Communication Disorders' Research Facility Unit at the University of Canterbury. Participants performed all experiments on the same day. Participants in the study reported in Chapters 3 and 4 completed three tasks: a reading comprehension task, a probe-paradigm task and an assessment of working memory. Participants in the study detailed in Chapter 5 completed five tasks: four comprehension experiments and an assessment of working memory. All participants performed the experiments in a set order—comprehension experiments followed by the working memory capacity assessment. All experiments were automated and presented by a computer. The tasks were presented on a 22-inch widescreen LCD monitor. Instructions were identical across participants and provided in written text on the computer screen.

The comprehension experiments were modelled on Gernsbacher's (1989) probe-response paradigm and consisted of two subtasks: a reading comprehension task and a probe recognition task. Previous studies (Carreiras et al., 1995; Gernsbacher, 1989, 1997b; Gernsbacher & Hargreaves, 1988; Gernsbacher et al., 1989; Gernsbacher & Jescheniak, 1995; Gernsbacher et al., 2001; Kim et al., 2004; Noh & Stine-Morrow, 2009) have demonstrated that the probe-response paradigm allows for the examination of offline as well

as online discourse comprehension. The comprehension task provides offline measures of comprehension accuracy. The probe recognition task enables the study of the processes of laying the foundation for the mental representation of discourse and suppression process during discourse comprehension through measuring the accessibility of the first and second mentioned names, and the relevant and irrelevant concepts. Using this paradigm, studies have provided evidence for the assignment of referents to anaphoric pronouns by measuring the responses to the referent and nonreferent names before and after the pronouns. Anaphora resolution has been demonstrated by the suppression of the nonreferents (Gernsbacher, 1989; Love & McKoon, 2011).

Moreover, the text presentation method used by Gernsbacher (1989) was suitable for the purpose of this study. Gernsbacher used a Rapid Serial Visual Presentation (RSVP) method in which text was presented word by word (presentation time = 300 ms + 16.667 ms per letter) in the centre of the screen. This method has been proved to improve comprehension performance in short passages when presentation rate was equal to, or longer than, 250 ms per word (Rayner & Clifton, 2002). Moreover, it prevents the readers from going back to search for names and therefor removes the effect of repetition on accessibility of names. Although comprehension in RSVP paradigms requires higher cognitive processing (D. Lee & Newman, 2010), language users are experienced in this type of language processing as it is close to the sequential nature of the natural conversational and spoken speech. Comprehenders of spoken speech are not capable of controlling the speech rate or to play back the earlier words and phrases (Susan & David, 1992).

2.3.1 Automated Reading Span Task

The Automated Reading Span Task (Unsworth, Heitz, Schrock, & Engle, 2005) was used to assess working memory capacity. Participants were presented with sets of sentences,

and each set included between two and seven sentences. Some of the sentences were not meaningful (e.g., *the prospector's dish was lost because it was not based on fact*). Participants' first task was to judge whether the sentence made sense or not. After each sentence they were presented with the sentence *"this sentence makes sense"* and two choices; TRUE or FALSE. They were to answer by clicking the mouse on the TRUE or FALSE buttons on the screen. After the participant entered their response, a letter (B, F, H, J, L, M, Q, R, or X) appeared on the screen. The participants' second task was to remember the letters. At the end of each set, participants were asked to recall all the letters in order. Two span scores were measured for each participant. First was the sum of the letters in all the sets that were fully correctly recalled. The second score was the sum of the all the correctly recalled letters in the correct order across all sets. The scores used for analysis in our study were the second score which was the total number of correctly recalled letters (see Unsworth et al., 2005 for details of experiment design, stimulus and scoring). An example of a two-sentence set is as follows:

Example:

Sentence 1: the prospector's dish was lost because it was not based on fact

This sentence makes sense *TRUE* *FALSE* Answer: FALSE

L

Sentence 2: during the winter you can get a room at the beach for a very low rate.

This sentence makes sense *TRUE* *FALSE* Answer: TRUE

X

RECALL: L, X

2.3.2 Comprehension Experiment

In the study detailed in Chapters 3 and 4, each participant performed two comprehension experiments with a break between. One experiment had lower working memory demand (Experiment 1) and the other had higher working memory demand (Experiment 2). The experiments were counterbalanced across participants, such that half of the participants completed Experiment 1 first (detailed in chapter 3), and the remaining participants completed Experiment 2 first (detailed in chapter 4). In the research reported in Chapter 5, each participant performed four comprehension experiments (Experiments 3, 4, 5, and 6) with a break after each experiment. The first two experiments (Experiment 3, and Experiment 4) in the Chapter 5 were short versions of Experiments 1 and 2 in Chapters 3 and 4 respectively. Details of the experiments are provided in the relevant chapter.

All comprehension experiments required participants to read a series of sentences on the computer monitor and answer the questions that followed. As they were reading the sentences, they were also periodically tested with a probe recognition task. Each trial consisted of a clause that introduced two characters one of which was referred to by an anaphoric pronoun later in the text (full details of the experimental stimuli in each experiment are provided in a subsequent section). Ten practice trials were undertaken before each experiment commenced. The first two trials practiced the comprehension task (e.g. Judy made Ruth a sandwich for breakfast and she toasted the bread. QUESTION: Who toasted the bread to make sandwich?) and the next four trials practiced the probe recognition task. The last four trials involved practice of both tasks simultaneously, which emulated the actual experiments. Participants used only one hand (the hand they normally wrote with) to press the response keys throughout the experiment. They were instructed to use their pointer finger to press one key, and their middle finger to press the other key. Information about the

experimental stimuli is provided in the next section, followed by specifics of the experimental procedure.

2.3.3 Sentence Stimuli

The experimental stimuli in all experiments were English sentences adapted from those of Gernsbacher (1989). 112 single sentences were used in Experiment 1 (Chapter 3) and 72 four-sentence paragraphs in Experiment 2 (Chapter 4). In Chapter 5, 50 single sentences were used in Experiment 3, 36 four-sentence paragraphs in each of the Experiments 4 and 5, and 36 sentence pairs in Experiment 6. The experimental stimuli in each experiment included a number of "lure" sentences/paragraphs in addition to experimental sentences/paragraphs. Names used in the probe task were not used in the lure sentences/paragraph. The specifics of the stimuli used in each experiment are detailed in relevant chapters. To ensure that the stimuli were culturally appropriate to New Zealand participants, four native speakers of English (two of them were native speakers of NZ English), naïve to the experiments reviewed the sentences. Culture-dependant sentences were adapted to ensure appropriateness for NZ English participants.

In all experiments, the first main clause contained two male or female characters, termed N1 (first person mentioned in the clause) and N2 (second person mentioned in the clause). One of the characters was referred to in the last clause by an anaphoric pronoun (either "he" or "she"). Example sentence are as follows:

Example 1: Bill (N1) handed John (N2) some tickets to a concert but he (anaphora) took the tickets back immediately

Example 2: Andy (N1) tried to beat Gary (N2) in a game of chess but he (anaphora) managed to win every time.

The discourse characters all exhibited relatively common names in English. The names chosen for the experimental tasks were all among the first 150 most popular English names for male and female babies born during 1970 – 1979 as determined by the official website of the US social security administration (2013). The names, which were uncommon in New Zealand (e.g. Jose), and those that had a pronunciation similar to other names were excluded from the list. An equal numbers of male and female characters were presented in the stimulus lists. Names and anaphoric pronouns were all gender consistent within a trial. They were also matched according to their number of characters and familiarity.

Stimulus sentences were constructed so that there were five words between N2 and the end of the first clause in Experiments 1 and 3, and between N2 and the end of the first sentence in Experiments 2, 4, 5, and 6. Moreover, there were always five words between the anaphoric pronoun and the end of the sentence. Additionally, to ensure that the stimuli were culturally appropriate to New Zealand participants, four native speakers of English (two of them were native speakers of NZ English), naïve to the experiments reviewed the sentences. Culture-dependant sentences were adapted to ensure appropriateness for NZ English participants.

In half of the experimental sentences in each experiment, N1 was the reference of the anaphoric pronoun (as in Example 1) and in the other half N2 was the reference (as in Example 2). Gernsbacher (1989) conducted a normative study on the sentences to ensure that in each sentence only one of the names was favourably biased by the sentence context. This insured that the anaphora in each sentence had a unique referent. In Experiments 2, 4, 5, and 6, lure sentences used the same structure as the experimental sentences (half with N1 as anaphora referent and half with N2 as anaphora referent). In Experiment 1, some lure

sentences differed in that their anaphoric pronoun was the plural pronoun *they* instead of *he/she* (e.g. Wendy read Stacy a novel about a couple and they were saddened by the story).

2.3.4 Experimental Probes for Recognition Tasks

The probes presented in the recognition tasks in all experiments were selected using the same procedure as in selecting the names for N1 and N2. Half of the names used in the tasks were male and the other half were female names. The probe used for each of the experimental sentences/paragraphs, was one of the names mentioned in the sentence/paragraph (first/second). The probes for the lure sentences/paragraphs were all the names that did not occur in the sentences/paragraphs (FOIL Probe) but were matched with the sentence/paragraph names according to the number of characters, gender and familiarity. For half of the comprehension trials, the probe names were the referent of the anaphora (REF), and for the other half the nonreferent (NREF). In addition, half of the names in each condition were the first mentioned names in the sentence (N1) and half were the second mentioned names (N2). Response times and accuracy of responding to the probes were measured.

2.3.5 Comprehension Questions

Following the presentation of each sentence in Experiments 1 and 3, and each paragraph in other experiments, participants were presented with a forced-choice comprehension question. The questions asked about the referent of the pronoun in the previous sentence or paragraph. The answers to these questions were N1 in half of the cases and N2 in the other half. The questions were always about the last clause asking about the referent of the anaphora (Example 1 below). The questions for the lure sentences with the plural anaphoric pronoun, “they”, used in Experiment 1, were always about the first clause (Example 2 below).

Example 1:

Sentence: Roy walked Joe over to the dentist's office but he waited outside in the lobby.

Question: Who waited outside in the lobby? Answer: Roy

Example 2:

Jodi fixed Beth up on a blind date and they went shopping for the date.

Question: Who had a blind date? Answer: Beth

2.3.6 Experimental procedure for comprehension tasks

Each comprehension experiment consisted of a comprehension task and a probe recognition task. Participants were presented with a series of sentences followed by a comprehension question. Following Gernsbacher's (1989) paradigm, each trial began with a centred "+" sign for 750 ms, which was replaced by the sentence(s) presented one word at a time. Word presentation time was dependent upon the length of the word. Each word was presented for 300 milliseconds plus 16.667 milliseconds per letter. Each word was followed by an inter-word interval of 150 ms.

Trials were displayed in a randomized order. Also for each participant, the names in NAME 1 and NAME 2 positions were randomized. Following each sentence/paragraph, the word "Test" appeared on the screen and was displayed for 750 milliseconds towards the bottom of the screen. Then the comprehension question and two answer choices were displayed. One of the choices was presented on the bottom left corner and the other on the bottom right corner of the screen. The answer choices in each corner were correct half of the time. Participants were instructed to use one of the two arrow keys on the keyboard to answer

each question. They were asked to press the key on the LEFT to indicate that the answer that appeared on the left side of the screen was correct and to press the key on the RIGHT to indicate that the answer that appeared on the right side of the screen was correct. The question and choices remained on the screen until a response was entered or until 10 seconds passed. After entering each response, participants received accuracy feedback (Correct, Incorrect, or No Response).

2.3.7 Experimental procedure for probe recognition task

Probe names were displayed in capital letters at the top of the screen. When a probe name appeared on the screen, participants were to decide whether that name had occurred in the text they were reading. Participants answered the recognition task by pressing the same keys they used to answer the comprehension questions. Those keys on the keyboard (right and left arrow keys) were labelled “YES” and “NO”. They also had “right” and “left” arrows on them. Participants were instructed to press the key labelled “YES” if the word had already appeared in the text and press the key labelled “NO” if it had not. They were asked to answer with the same hand that they answered comprehension questions with. Probe names remained on the screen until a response was entered or until 2.5 seconds passed. Probes were displayed at two different testing points. In half of the trials in each condition probes were presented before the anaphora and in the other half after the anaphora. Details of each testing point in each experiment are provided in the relevant chapters.

2.4 DATA AND STATISTICAL ANALYSIS

Two dependant variables were employed: (a) accuracy and (b) response time to the probe recognition task. For (a) two accuracy scores were calculated for each participant in each experiment—both accuracy in response to the comprehension question and accuracy in completion of the probe recognition task.

A participant's data was removed if the participant completely or partially failed to perform the recognition task, or had outlier comprehension accuracy scores (below 2SD from the mean of all the participants' mean accuracies) in any of the comprehension experiments, or if he was unable to perform the reading span task. This exclusion was necessary to make sure that the response times reflected the levels of accessibility of the probes and were not based on chance responses. In addition, the exclusion minimized the effects of chance responses to comprehension questions, making the probe response-times a meaningful measure of online comprehension processing.

In total, data from 27 out of 30 younger and 20 out of 30 older participants were included in the full analysis for Chapters 3 and 4. All participants had above-chance accuracy in probe recognition task. Data from 35 out of 40 participants were included for Chapter 5. For analysis of the RTs to the probe recognition task, only experimental trials that were responded to correctly (in both probe recognition and comprehension accuracy tasks) and with raw RTs greater than 300 milliseconds were included.

A cutoff point of 300 milliseconds was chosen based on the belief that genuine response times cannot be shorter than 100 milliseconds which is the time required for physiological processes. Moreover, it has been suggested that response times faster than 300 milliseconds are too short to allow for a conscious response. Therefore, response times shorter than 300 milliseconds are attributed to unintentional responses (Hermans, De Houwer, & Eelen, 2001; Hermans, Spruyt, & Eelen, 2003; Whelan, 2008).

The remaining data were then further trimmed to exclude raw RTs larger than 2SD from the mean of all the correct experimental trials in each experiment. A more conservative data trimming procedure (Van Selst & Jolicoeur, 1994) was also tried. However, the results were not affected by the kind of trimming procedure. Since raw RT distribution is usually

positively skewed, suitability of normal, logarithmic and inverse Gaussian distributions were evaluated using the Shapiro Wilks test. The test results suggested that logarithmic distribution best suited the data. Therefore, logarithmic transformation was used to normalize the RT distribution. Mixed effects modelling (MEM) was used to analyse comprehension accuracy and response times. Recently, MEMs have been preferred over traditional methods such as ANOVAs, as MEMs are able to account for the individual variances (e.g. different characteristics of participants and stimuli) (Baayen, Davidson, & Bates, 2008; Baayen & Milin, 2010)

3. CHAPTER THREE

Age-related Changes in Anaphora Resolution

3.1 ABSTRACT

This study investigated how ageing and working memory capacity affected anaphora resolution in a comprehension task. In addition, the influence of age and working memory on anaphora resolution was evaluated by examining the advantage of first mention and suppression of irrelevant information in a probe recognition task. Thirty younger and 30 older participants completed a reading comprehension task and probe recognition task, and an assessment of working memory capacity. Each sentence used in the reading comprehension task included an anaphoric pronoun, a referent and a nonreferent name. Participants' comprehension accuracy and response times were analysed.

Results of the study found that older adults were less accurate than younger adults in the comprehension of anaphoric pronouns. Comprehension accuracy was also affected by working memory capacity with better performance for those with higher working memory capacity, regardless of age. In the probe-recognition condition, working memory capacity, but not ageing, was related to advantage of first mention. Finally, neither age nor working memory was a factor in suppression of irrelevant information as measured by the probe-recognition tasks. In fact, when measured at the end of the sentence, response times to referent and nonreferent names were not different either in the younger or in the older participants. The findings of the study suggest that age and working memory capacity influence performance in anaphora resolution in a comprehension task. The findings of the probe-recognition task suggest that working memory capacity plays a significant role in laying the foundation of the discourse structure as measured by 'advantage of first mention'. The influence of age could not be determined in this study, as neither the older nor the younger participants appeared to suppress the irrelevant information. The findings suggest that further exploration is needed using higher-demand anaphora resolution tasks.

3.2 INTRODUCTION

As discussed in Chapter 1, ageing is usually accompanied by changes in cognitive abilities. Language comprehension is one cognitive ability that has been reported to decline with ageing (Au et al., 1989; Cohen, 1979; Daneman & Carpenter, 1980; Hancock et al., 2001; Kynette & Kemper, 1986; Nicholas et al., 1985; Shewan & Henderson, 1988; Ulatowska et al., 1986). Age-related comprehension decline has been mainly observed in discourse processing abilities, particularly in inferring the potentially ambiguous meanings of discourse such as *anaphora resolution* (Borella et al., 2011; Cohen, 1979; Hancock et al., 2001; Hannon & Daneman, 2009; Light & Capps, 1986; Ulatowska et al., 1986).

Anaphora resolution is an important area of discourse to study for a number of reasons. First, finding the referent of discourse anaphora is important as anaphora plays a crucial role in maintaining discourse's connectivity and coherence(Graesser et al., 1997). Second, anaphora resolution is consistently noted to undergo changes with ageing (anaphora resolution has been defined and described in detail in Chapter 1, Section 1.5.1) (Cohen, 1979; Light & Capps, 1986; Ulatowska et al., 1986).

When considering the changes in anaphora resolution that accompany ageing, a decline in the capacity for maintaining information in working memory and an inability to suppress irrelevant discourse information are among the proposed causes (Borella et al., 2011; Brébion, 2003; De Beni et al., 2007; Hannon & Daneman, 2009; Hasher & Zacks, 1988; Light & Capps, 1986) Structure Building Framework (Gernsbacher, 1990) is one model of discourse comprehension that acknowledges the contribution of the maintenance and suppression functions of working memory and is a useful framework for exploring the underlying mechanisms of anaphora resolution.

As noted in Chapter 1, SBF suggests that language is processed through building mental representations for discourse. Early information in discourse constitutes the foundation of the mental representation while incoming information is mapped onto it. For example, in a discourse that introduces more than one character, the first character forms the foundation while the second character would be added to the already-laid foundation of the mental representation. Being part of the foundation, the first character needs to be enhanced and maintained activated. As a result, first mentioned character would be more accessible than the second mentioned character in readers' memory. This phenomenon has been referred to as "the advantage of first mention" and is a useful measure of working memory storage and maintenance. The next step of discourse processing, the updating of the information, is, according to SBF, resolved by the processes of suppression and enhancement when new information is received. When new information is received, it affects the accessibility of the previously stored information. Relevant information is enhanced by increasing its activation while irrelevant information is suppressed by decreasing its activation. The result is that relevant information should be accessed more quickly. This can be measured by on-line tasks that look at reaction time in response to probes that compare relevant and irrelevant information.

In SBF, storage and maintenance of information (laying the foundation) and suppression have been measured using a probe-response paradigm. Using this paradigm studies have documented the processes of laying the foundation through comparing the accessibility of the first mentioned and the second mentioned discourse concepts. The process of suppression has also been examined through comparing the accessibility of the relevant and the irrelevant discourse concepts (Carreiras et al., 1995; Gernsbacher, 1989; Gernsbacher & Hargreaves, 1988; Gernsbacher et al., 1989). The concepts of interest in a probe-response paradigm (e.g. relevant vs. irrelevant concepts, first mentioned vs. second mentioned

discourse characters) were presented as a probe to the participants while they were reading a passage. Participants were asked to decide whether they had seen the probe in the passage they were reading. The faster that participants recognised a previously-seen probe, the more accessible the probe was in their memory (Carreiras et al., 1995; Gernsbacher, 1989; Gernsbacher & Hargreaves, 1988; Gernsbacher et al., 1989). The processes involved in discourse comprehension in SBF are mainly the functions of working memory. Therefore it might be expected that differences in working memory abilities could affect discourse comprehension.

The cognitive processes involved in anaphora resolution have been explained in SBF. Moreover, two important processes of discourse comprehension in which working memory plays crucial role, can be examined in SBF. The early process of laying the foundation can be tested through examining the effect of age on the advantage of first mention of discourse characters. The first mentioned discourse character would be expected to be more accessible. Moreover, anaphora is only completely resolved when its referent is more accessible than the nonreferents. Studies have demonstrated that suppressing the accessibility of the nonreferent reflected the resolution of anaphora (Gernsbacher, 1989; Love & McKoon, 2011). Therefore, investigating how individuals' ages affect suppression of the nonreferent during anaphora resolution can provide an online assessment of age-related changes in anaphora resolution.

Individual differences in working memory capacity might also be expected to affect the speed at which anaphora is resolved. People with lower working memory capacity might be expected to resolve anaphora faster so that irrelevant information is suppressed earlier and more capacity can be devoted to processing the remaining of the sentence. Similarly, if older adults' difficulties in anaphora resolution is caused by a decline in working memory, they might be expected to suppress the nonreferent earlier due to smaller capacity for maintaining

information. In contrast, the inhibitory decline theory of ageing might predict an inability to suppress anaphora's nonreferents during anaphora resolution in older adults.

In summary, although working memory capacity and executive functions are believed to play important roles in anaphora resolution, the contributions of individual differences in working memory to age-related changes in anaphora resolution has remained unclear. Studies so far have demonstrated a decline in anaphora resolution accuracy and in working memory capacity in older adults. However, it is still unclear whether changes in the process of anaphora resolution are due to age-related differences in working memory capacity, or other age-related cognitive deficits. Knowledge of discourse comprehension difficulties in ageing and its underlying causes is important in maintaining and improving communication with older adults. In SBF, it is possible to examine age-related changes in two important sub-processes involved in anaphora resolution; laying the foundation for comprehension as demonstrated through the advantage of first mention, and suppression of nonreferents in the process of information update. Since working memory functions play crucial roles in these processes, the contributions of differences in working memory abilities to any change in the process of anaphora resolution in ageing can also be investigated.

3.2.1 Current study

The current study aimed to further investigate when, in the process of anaphora resolution, the change of accessibility of nonreferent and referent names occurred in younger and older adults. It also examined whether the accuracy and process of anaphora resolution were affected by individuals' age and working memory capacity. The dependant variables included participants' accuracy in answering anaphoric comprehension questions and response times in a probe recognition task.

Response times were used to examine 1) how accessible the first mentioned names in the discourse were compared to the second mentioned ones, and 2) how accessible the anaphora's referents were compared to nonreferents. The accessibility of the discourse entities reflected their levels of activation in working memory. The differences in accessibility caused by the order of mention (i.e. the advantage of first mention) were to reflect the process of laying the foundation during discourse comprehension. The accessibility of referent and nonreferent names was measured at different points, while participants were reading anaphoric sentences, to investigate the change of accessibility of referent/nonreferent names during anaphora resolution. The change of accessibility of referent and nonreferent names was to demonstrate the suppression process. The time-course for the suppression of the nonreferents reflected when the process of anaphora resolution was completed.

Application of Gernsbacher's (1989) paradigm allowed for the examination of the hypothesis that age-related changes in anaphora resolution might be due to inefficient suppression of the unrelated information. Inefficient suppression has been attributed to an inhibitory deficit (Hasher & Zacks, 1988). Therefore, the inhibitory decline theory of cognitive ageing could be partially supported if older adults differ from younger adults in suppressing the nonreferents. As discussed in Chapter 1, Section 1.2.2.2, studies have provided evidence that not all inhibitory functions decline with ageing. Inhibitory functions are suggested to be goal-related and different inhibitory processing might be used for different task goals (Lustig et al., 2007). Therefore, examining older adults' suppression abilities during anaphora resolution would provide an insight into age-related changes in comprehension-related inhibitory functions.

Three primary questions were addressed: 1) Do age and working memory capacity affect anaphora resolution in a comprehension task?; 2) Do age and working memory affect advantage of first mention in a probe recognition task (as measured by interactions among age, probe-type, and referent position, and working memory capacity, probe-type, and referent position)?; and, 3) Does age affect suppression of irrelevant information in an anaphora resolution task (as measured by the interaction between probe type and probe testing point)?

It was hypothesized that age and working memory would affect comprehension with: a) older adults being less accurate than younger adults in comprehension of anaphoric pronouns and b) higher working memory being associated with better comprehension. It was further predicted that working memory capacity would affect the advantage of first mention. First mentioned names would be expected to be more accessible than the second mentioned names in people with higher working memory capacity compared to those with lower working memory capacity because working memory capacity plays crucial role in maintaining activated information accessible. Given that older individuals might be expected to have lower working-memory capacity, age would also be expected to be related to the advantage of first mention.

With regard to suppression, it was expected that age would have an influence on the suppression of irrelevant information. Older adults would be expected to suppress the irrelevant information (i.e. nonreferents) earlier if age-related changes in anaphora resolution are due to working memory capacity decline. However, if changes in anaphora resolution in older adults were due to inefficient suppression, they would be expected not to suppress the nonreferents. in such case, older adults would not be as capable as younger adults in suppressing the accessibility of the nonreferent names.

3.3 METHOD

The study received ethical approval from the University of Canterbury Human Ethics Committee. All individuals provided written consent to participate.

3.3.1 Participants

Participants included 30 younger (6 males and 24 females, *M = 25.33, SD = 4.31*, range = 19-35 years) and 30 older (14 males and 16 females, *M = 74.15, SD = 4.64*, range= 66-87 years) right-handed native speakers of New Zealand English (NZE). They reported no history of neurological disease, dementia, cognitive impairment, cardiovascular disease, uncontrolled hypertension, learning disability, attention deficit disorder, or speech disorder.

3.3.2 Procedure

Participants in the study completed two tasks: a comprehension experiment (Experiment 1) and an assessment of working memory—the Automated Reading Span Task (Unsworth et al., 2005). The comprehension experiment was modelled on Gernsbacher (1989) and consisted of two subtasks: a reading comprehension task and a probe recognition task.

3.3.2.1 Experiment 1

The reading comprehension experiment required participants to read a series of sentences on the computer monitor and answer the questions that followed. As they were reading the sentences, they were also periodically tested with a probe recognition task. Each comprehension sentence consisted of two clauses; the first clause introduced two characters one of which was referred to by an anaphoric pronoun in the second clause (see Chapter 2, Section 2.3.2 for details of the experimental design). Details of the sentence stimuli are

included below, followed by specifics of the experimental procedure. Information about the experimental probes and comprehension questions are detailed in Chapter 2, Sections 2.3.4 and 2.3.5)

3.3.2.2 Sentence Stimuli

The experimental stimuli were 112 English sentences adapted from those of Gernsbacher (1989)—64 were experimental sentences (See Appendix A), with the remaining 48 "lure" sentences. The current study only investigated the resolution of pronoun anaphora. Therefore, the noun anaphora of Gernsbacher's sentences was substituted with pronoun anaphora. Removing the effect of anaphora type from the study design provided us with a doubled number of trials in each condition.

Each sentence contained two main clauses connected by a conjunction. The first clause contained two male or female characters, termed N1 (first person mentioned in the sentence) and N2 (second person mentioned in the sentence). One of the characters was referred to in the second clause by an anaphoric pronoun (either "he" or "she"). An example sentence is as follows:

Example 1: Bill (N1) handed John (N2) some tickets to a concert but he (anaphora) took the tickets back immediately

Example 2: Andy (N1) tried to beat Gary (N2) in a game of chess but he (anaphora) managed to win every time.

The second set of 48 sentences—all "lure"— were also adapted from Gernsbacher's (1989) study. Names used in the probe task were not used in the stimuli sentences. Of these 48 lure sentences, 32 used the same structure as the 64 experimental sentences (16 with N1 as anaphora referent and 16 with N2 as anaphora referent). The remaining 16 sentences differed

in that their anaphoric pronoun was the plural pronoun *they* instead of *he/she* (e.g. Wendy read Stacy a novel about a couple and they were saddened by the story).

3.3.2.3 Experimental Procedure for Experiment 1

The comprehension experiment consisted of a comprehension task and a probe recognition task. Experimental procedures for the comprehension task are detailed in Chapter 2, Section 2.3.6. In brief, participants were presented with a series of sentences followed by a comprehension question. While reading the sentences, participants were presented with a probe. General information about the probe recognition task is provided in Chapter 2, Section 2.3.7. Probes were displayed at two different testing points. In half of the trials in each condition, probes were presented before the anaphora and in the other half after the anaphora. Details of each testing point are as follows:

1. Before-anaphora testing point: Probes were presented 150 milliseconds after the offset of the last word of the first clause (e.g. Bill handed John some tickets to a concert *PROBE* but he took the tickets back immediately).

2. After-anaphora testing point: Probes were presented 150 milliseconds after the offset of the last word of the second clause. In other words, they were given at the end of the sentence (e.g., Bill handed John some tickets to a concert but he took the tickets back immediately. *PROBE*).

3.3.3 Data and Statistical Analysis

Two dependant variables were employed: (a) accuracy and (b) response time to the probe recognition task. For (a), two accuracy scores were calculated for each participant—both accuracy in response to the comprehension question and accuracy in completion of the probe recognition task. The data from 12 participants (three younger and nine older) were

removed as they completely or partially failed to perform the recognition task, or had outlier comprehension accuracy scores (below 2SD from the mean of all the participants' mean accuracies) (see Chapter 2, Section 2.4 for details of participant exclusion). Data from one older participant was excluded, as he was unable to perform the reading span task. The remaining participants' working memory spans ranged from 36 to 75 in the younger group (*Mean = 63.3, Median = 65.0, SD = 9.15*) and 32 to 75 in the older group (*Mean = 53.80, Median = 53.50, SD = 10.93*).

In total, data from 27 out of 30 younger and 20 out of 30 older participants were included in the full analysis. Statistical information about participants' mean scores in comprehension and probe recognition tasks in Experiment 1 is provided in Table 3.1.

Table 3.1: *Participants' mean accuracies in the comprehension and probe recognition tasks*

(%) Accuracy	Group	Max	Min	Median	Mean	SD
Comprehension	Older	94.79	52.08	76.56	76.04	0.11
	Younger	98.96	73.96	92.71	90.74	0.06
Probe Recognition	Older	100	68.75	90.62	89.01	0.08
	Younger	100	88.54	95.83	95.76	0.03

For analysis of the RTs to the probe recognition task, only experimental trials that were responded to correctly (in both probe recognition and comprehension accuracy tasks)

and with raw RTs greater than 300 milliseconds were included. The remaining data were then further trimmed to exclude raw RTs larger than 2SD from the mean of all the correct experimental trials (1896 ms). These cut-offs removed 4.3% of the data. Data trimming procedures were required to ensure that only meaningful response times that reflected the process of correct anaphora resolution were analysed. Removing the trials with incorrect comprehension accuracy excluded the possibility that response time patterns were attributed to wrong resolution of anaphora due to either information loss or switching the roles and the order of mention of the first and the second characters in mind. Logarithmic transformation was used to normalize the RT distribution. Mixed effects modelling (MEM) was used to analyse comprehension accuracy and response times.

3.4 RESULTS

3.4.1 Comprehension Accuracy

A series of binomial mixed effects models were used to analyse participants' accuracy of responses to the comprehension questions. The analysis examined the fixed effects of trial number, age group (younger or older), genders, years of education, gender of the sentence characters, working memory capacity (Reading Span score), position of the referent in the sentence (order of mention-first or second), testing point for the recognition task (before reading the anaphora or at the end of the sentence), log-transformed RT of the probe recognition task in the preceding trial, and the accuracy of response to both the preceding comprehension question and probe recognition task. All relevant interactions between these fixed effects were also tested. Random effects for participant and stimuli were also included. Including the random slopes in this model did not affect the significant results.

Model fitting was performed in a backward-stepwise fashion, followed by forward fitting of maximal random effects structure. Models were evaluated by model fitness

comparisons using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The final model for comprehension accuracy is presented in Table 3.2. Random effects included in the final model for comprehension accuracy are presented in Table 3.3.

Table 3.2: *Coefficients and p values of the binomial mixed effects model for comprehension accuracy in Experiment 1, with participant and stimulus as random effects.*

Effects	**Coefficient**	**Standard Error**	**z value**	**P value**
(Intercept)	-0.503	0.521	-0.964	0.335
Trial number	0.149	0.051	2.932	0.003**
RP: Second	0.341	0.162	2.103	0.035*
PTP: END	-0.353	0.162	-2.177	0.029*
Age group: younger	0.862	0.198	4.358	< 0.001**
WM Capacity	0.027	0.009	2.934	0.003**
Accuracy in PCT	0.490	0.132	3.699	< 0.001***

*Note: WM = working memory, PTP = probe testing point, PCT= preceding comprehension trial, RP: referent position, END = end of the sentence, * = p < 0.05, ** = p < 0.01, *** = p < 0.001*

Table 3.3: *Random effects included in the final binomial mixed effects model for comprehension accuracy in Experiment 1.*

Random effects		**Variance**	**Standard Deviation**
Stimulus	(Intercept)	0.223	0.479
Participant	(Intercept)	0.219	0.468

The model revealed a significant main effect of age group on comprehension accuracy *($p < 0.001$)*, indicating that comprehension accuracy was significantly higher in the younger group compared to the older group. The main effect of working memory capacity was also significant *($p < 0.01$)*, demonstrating that as working memory capacity increased, so did the comprehension accuracy.

The main effect of referent position in the sentence also reached significance ($p < 0.05$), showing that participants completed the task with greater accuracy when the referent was the second mentioned name in the sentence. As expected, performance improved across the experiment, with a significant main effect of trial number ($p < 0.01$). Moreover, testing point for the probe recognition task ($p < 0.05$) and performance on the preceding comprehension task ($p < 0.001$) also significantly affected performance on the comprehension task. Comprehension accuracy was lower in trials in which the probe recognition task was presented at the end of the sentence. Furthermore, participants were more accurate in answering comprehension questions after correctly answering the preceding comprehension trial.

3.4.2 Response Times

A series of linear mixed effects models were used to analyze participants' response times (log-transformed RTs) to the probe recognition task. The analysis examined the fixed effects of age group (younger or older), participants' genders, participants' years of education, gender of the sentence characters, participants' working memory capacity (Reading Span score), probe type (referent (REF) or nonreferent (NREF)), probe testing point (before anaphora or at the end of the sentence), referent position in the sentences (first or second), probe position in the sentences, trial number, log-transformed RT to the probe recognition task in the preceding trial, and the accuracy of responses to both the preceding

comprehension question and probe recognition task. All relevant interactions between fixed effects were tested. The interactions that were to reflect the two underlying processes during anaphora resolution were 1) the interaction of probe type and probe testing point which was to demonstrate the process of suppression, 2) the interaction of probe type and referent position in the sentence which was used to examine the advantage of first mention and thus reflected the process of laying the foundation. Random effects for participant and stimulus were also included.

Model fitting was performed in a backward-stepwise iterative fashion, followed by forward fitting of maximal random effects structure. Model fitting was independently supported by model fitness comparisons using AIC and BIC. It should be noted that some results did not hold when the fullest random effects structure was included. The final model for log-transformed RTs is presented in Table 3.4.

Table 3.4: *Coefficients and p values of the final linear mixed effects model for RTs in Experiment 1, including fullest random effects structure.*

Effects	**Coefficient**	**Standard Error**	**t value**	**P value**
(Intercept)	7.160	0.118	60.88	
Trial Number	-0.042	0.007	-6.38	< 0.001***
Age Group: Younger	-0.280	0.040	-6.96	< 0.001***
RP: Second	0.105	0.070	1.49	0.136
RT in PRT	0.001	0.001	5.47	< 0.001***
Accuracy in PRT	-0.046	0.017	-2.68	0.007**
WM Capacity	-0.001	0.002	-0.34	0.734
Probe Type: REF	0.136	0.068	2.01	0.044*

WM Capacity*RP: Second	-0.002	0.001	-1.78	0.075
Probe Type: REF*RP: Second	-0.166	0.098	-1.69	0.091
WM Capacity*Probe Type: REF	-0.003	0.001	-2.4	0.016*
WM Capacity*Probe Type: REF*RP: Second	0.003	0.002	2.02	0.043*

*Note: WM = working memory, RP = referent position, PRT = preceding recognition trial, RT = response time, REF = referent, * = p < 0.05, ** = p < 0.01, *** = p < 0.001*

This model included random intercepts for stimuli and participants, as well as by-participant random slopes for trial number, probe type, referent position, both RT and the accuracy of response to the preceding probe recognition task, and the interaction of probe type and referent position. Including these random effects allowed the model to take into account the mean differences in RTs across stimuli and across participants, as well as variable sensitivity to the effects of task features across participants. The random intercepts and by-participant random slopes included in the final model are presented in Table 3.5.

Table 3.5: *Random effects and slopes included in the final linear mixed effects model for RTs in Experiment 1.*

Random effects		**Variance**	**Standard Deviation**
Stimulus	(Intercept)	<0.001	0.017
Participant	(Intercept)	0.026	0.162
	Trial Number	0.001	0.038
	Probe Type : REF	<0.001	0.021
	RP: Second	<0.001	0.031
	RT in PRT	<0.001	<0.001
	Accuracy in PRT	0.001	0.032

	Probe Type: REF*RP: Second	0.002	0.044
Residual		0.033	0.181

Note: RP = referent position, PRT = preceding recognition trial, RT = response time, REF = referent

Question 2 aimed to determine whether age and working memory capacity affect advantage of first mention in a probe recognition task. Relevant interactions testing these questions included that between working memory, referent type and referent position in the sentence, and that between age, referent type and referent position. Statistical analysis revealed significant interaction of working memory capacity, probe type and referent position *($p < 0.05$)*. This three-way interaction is depicted in Figure 3.1. Participants with larger working memory capacity were faster at responding to the probes when the probes were the first mentioned names in the sentence compared to when they were the second mentioned names. In contrast, participants with lower working memory capacity were faster at responding to the probes when they were the second mentioned names in the sentence. The interaction between age, referent type and referent position, however, was not significant and was removed from the model in the process of model development.

Additional fixed effects included in the model that remained significant were: a) main effect of age group *($p < 0.001$)* indicating that younger adults were faster in responding to the probes compared to the older adults, b) main effect of trial number *($p < 0.001$)*. As expected, participants' response time decreased across the experiment, c) main effect of Probe type *($p < 0.05$)* demonstrating that participants were generally faster in responding to a probe when the probe was the referent of the pronoun compared to when it was not the referent of the pronoun, d) main effect of response time to the preceding probe recognition trial *($p < 0.001$)* showing that participants were slower in answering a probe if they had a higher response time in the preceding probe recognition trial, e) main effect of accuracy in the preceding probe

recognition trial *(p < 0.01)* demonstrating that participants were faster in responding to a probe after correctly answering the preceding recognition trial, and f) the interaction of working memory capacity and probe type *(p < 0.05)* indicating that participants with a higher working memory capacity were significantly faster at responding to the probes when they were the referents of the pronouns.

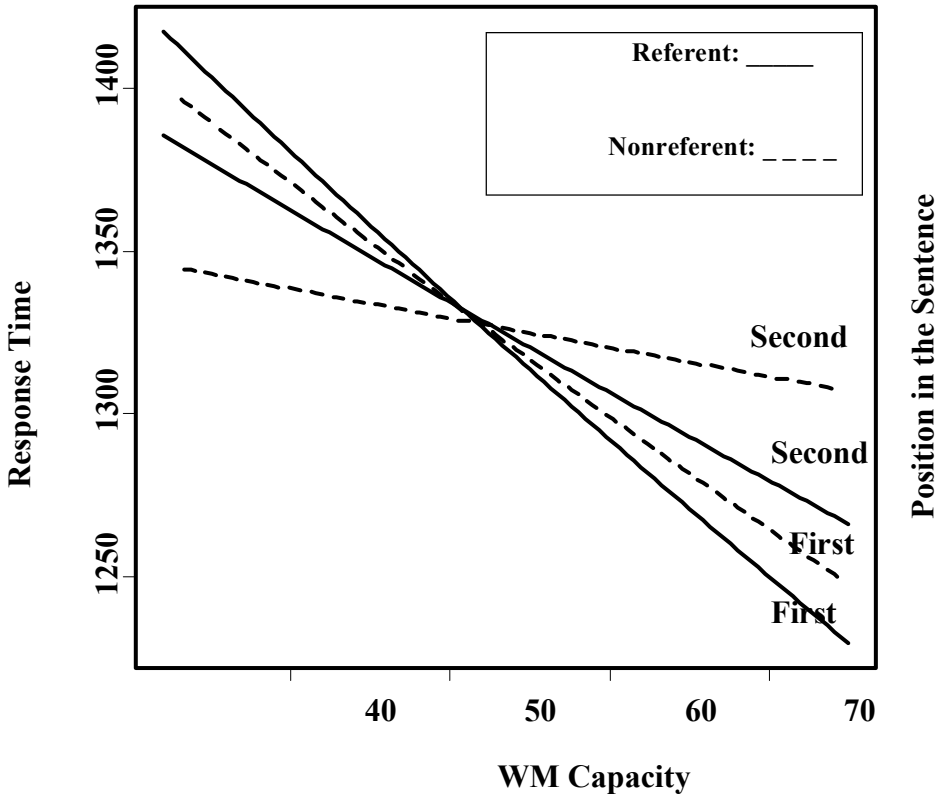

Figure 3.1: Three-way interaction of working memory capacity, probe type and referent position

The purpose of question 3 was to determine whether age affect suppression of nonreferent names during anaphora resolution. Relevant interaction testing this question was the interaction between age, probe type and probe testing point. This interaction was not significant and was thus removed from the model in the process of model development. In addition, the interaction of probe type and probe testing point which was to reflect the suppression process was also not significant. To make sure that the small effect size was not due to low statistical power, an imputation procedure was applied to the data through doubling and tripling the dataset. Increasing the statistical power, however, resulted in decreased effect size. Previous studies, which had found a significant interaction of probe type and probe testing point, had mainly used ANOVAs for statistical analysis of the RTs. While these analyses only included random intercepts for stimuli and participants, the final mixed effects model reported in this study included by-participant random slopes in addition. To examine whether any differences between the results obtained from the current study and the findings of the previous studies were due to the statistical techniques used for analysis, a separate model was developed to resemble a traditional ANOVA. The results from this analysis are provided in Table 3.6. This model revealed a significant interaction between probe type and probe recognition testing point *($p < 0.05$).*

Table 3.6: *Coefficients and p values of the linear mixed effects model for RTs in Experiment 1, including only the random effects of participants and stimuli.*

Effects	**Coefficient**	**Standard Error**	**t value**	**P value**
(Intercept)	7.15	0.121	59.14	
Trial Number	-0.042	0.007	-6.32	< 0.001***
PTP: END	0.036	0.016	2.2	0.028*

Age Group: Younger	-0.277	0.040	-6.91	< 0.001***
WM Capacity	<-0.001	0.002	-0.38	0.704
Probe Type: REF	0.149	0.073	2.03	0.042*
RP: Second	0.117	0.065	1.8	0.071
RT in PRT	<0.001	<0.001	5.4	< 0.001***
Accuracy in PRT	-0.045	0.017	-2.61	0.009**
PTP: END*Probe Type: REF	-0.035	0.016	-2.11	0.035*
PTP: END*RP: Second	-0.035	0.017	-2.11	0.035*
WM Capacity*Probe Type: REF	-0.003	0.001	-2.17	0.030*
WM Capacity*RP: Second	-0.002	0.001	-1.87	0.061
Probe Type: REF*RP: Second	-0.17	0.091	-1.88	0.060
WM Capacity*Probe Type: REF*RP: Second	0.003	0.002	2.25	0.024*

*Note: WM = working memory, RP = referent position, PTP = probe testing point, PRT = preceding recognition trial, RT = response time, REF = referent, END = end of the sentence, * = p < 0.05, ** = p < 0.01, *** = p < 0.001*

3.5 DISCUSSION

The current study investigated the influence of age and working memory capacity on anaphora resolution by examining 1) comprehension accuracy 2) advantage of first mention and 3) suppression of irrelevant information. First, the study found that both age group and working memory capacity affected comprehension accuracy. Second, the results showed that differences in working memory capacity affected the process of anaphora resolution by influencing the advantage of first mentioned name. Third, neither younger nor older adults suppressed the accessibility of the nonreferents by the time they finished reading the sentences. These primary results are discussed in separate sections.

3.5.1 Effects of Age and Working Memory Capacity on Comprehension Accuracy

The current study found that both age group and working memory capacity had a significant effect on the comprehension of anaphora. Firstly, the negative effect of age on the comprehension of short sentences is in contrast with the previous studies that have reported age effects on the comprehension of anaphora under high working memory storage load (long sentences), but not under lower storage load (Light & Capps, 1986). Additionally, working memory capacity affected comprehension of anaphora regardless of age—the effect held for both younger and older adults. Given that both age and working memory score affected the comprehension accuracy scores, it is not clear whether age-related decline in comprehension was caused by lower working memory capacity in older adults or was due to an age-related decline in other cognitive abilities.

The older participants had generally lower working memory capacities and lower comprehension scores. The older group had a lower median in reading span scores compared to the younger group. The differences in working memory capacity can explain the differences in comprehension accuracy in each age group. However, it is unclear whether the ageing effect can only be attributed to the differences in working memory capacity between older and younger groups. This issue can be further investigated by scrutinizing the underlying process of anaphora resolution using response times.

As mentioned earlier, it has been suggested that when facing ambiguity in a sentence, people with lower working memory capacity suppress the incorrect interpretation faster, which results in a single interpretation. That is because they do not have sufficient capacity for concurrent storing and processing of the remainder of the sentence. In contrast, high capacity readers might delay the process and maintain all interpretations until disambiguation is required as they have sufficient capacity to maintain all interpretations (Just & Carpenter, 1992). Older adults had lower working memory capacity. Therefore, if the age-related

difference in language comprehension were a result of a working memory capacity decline, it would be expected to observe an earlier suppression of the nonreferent name in older adults compared to the younger adults.

3.5.2 Effects of Age and Working Memory Capacity on the Advantage of First Mention

An aim of the current study was to investigate the effects of age and working memory capacity on the process of laying the foundation during anaphora resolution. This was addressed through examining how the advantage of first mention interacts with working memory capacity and age group. Current results suggest that individual differences in working memory capacity contributed to differences in the process of laying the foundation. Previous studies on anaphora resolution reported an advantage for the first mentioned compared to the second mentioned name in the sentence (Gernsbacher, 1989). Unlike what we expected, participants in this study were significantly more accurate in the comprehension of pronoun anaphora if the pronouns referred to the second mentioned name rather than the first mentioned name in the sentence. Considering the RTs in the probe recognition task, the findings of this study showed that the advantage of first mention only existed for participants with higher working memory capacity. In contrast, people with lower working memory capacity showed an advantage for the second mentioned name. The three-way interaction of working memory capacity, referent position and probe type revealed that participants with high working memory scores recognized the first mentioned name faster than the second mentioned name in the sentence, while participants with lower working memory scores showed the opposite RT pattern. This finding indicates that the most recent name had been the most activated and the most accessible discourse character for participants with lower

working memory capacity while first mentioned name had been more accessible for participants with higher working memory capacity.

As mentioned earlier, the advantage of first mention is considered to be the result of laying the foundation for mental representation in the process of comprehension (Gernsbacher, 1990). Higher accessibility of the second mentioned name in participants with lower working memory capacity suggests that individual differences in working memory capacity might affect the process of laying the foundation and maintaining the information that forms the foundation. In other words, the results of this study showed that participants with higher working memory capacity were better able to lay foundations for discourse comprehension. This finding is consistent with the SBF and the previous studies on the correlation of language comprehension and working memory capacity. SBF suggested that less skilled comprehenders with lower working memory capacity were less able to either lay foundations or update the information (Gernsbacher, 1990, 1997b; Perfetti, 1994). In contrast to the participants with high working memory capacity, those with lower working memory capacity relied more on the recency of the information instead of laying the foundation. This is also consistent with the results of a recent study by van Rij and colleagues (2013) on the influence of working memory capacity on reference resolution (van Rij, van Rijn, & Hendriks, 2013). In general, this finding suggests that individual differences in working memory capacity affected the process of laying the foundation for comprehension. Participants with higher working memory capacity were more accurate in comprehending pronoun anaphora, which might be due to their higher ability to maintain information as part of the foundation and thus their higher engagement in building the mental representation for comprehension. Unlike working memory capacity, age group did not affect the advantage of first mention.

3.5.3 Effects of Age on Suppression Process

A goal of the current study was to examine the effect of age on the process of suppressing anaphora's nonreferents. However, this study showed no significant suppression of nonreferents in either younger or older participants. The interaction between probe type and probe recognition testing point which was nonsignificant in the current study had been suggested by Gernsbacher (1989) to reflect the process of suppression. In her study, Gernsbacher (1989) found that RTs to nonreferent and referent names were not significantly different before reading the anaphora. After reading the anaphora, however, RTs to the nonreferent names increased while RTs to the referent names decreased significantly. It was suggested that referent became more accessible due to enhancement, whereas the competing nonreferent became less accessible due to suppression during the anaphora resolution process.

In contrast to Gernsbacher's (1989) finding, the interaction of probe type and probe recognition testing point did not reach significance in the current study. Although referents were significantly more accessible than nonreferents, there was no significant change of accessibility after reading the anaphora. A significant two way interaction also showed that while referents were more accessible for participants with higher working memory capacity, nonreferents were the most accessible names for participants with lower working memory capacity. Moreover, our results revealed no effect of age group on the interaction of probe type and testing point. In other words, our finding suggests that there was no suppression effect in either younger or older groups by the time participants finished reading the sentences.

It should be noted that statistical techniques might have some influence in the inconsistency between our results and Gernsbacher's. Gernsbacher used ANOVAs for

statistical analysis of the RTs. However, more recently, it has been suggested that MEMs can provide a better analysis of RTs, as they are better able to take into account the variations across stimuli and participants. The final model in the current study used maximal random effects structure, which, apart from random intercepts for stimuli and participants, included by-participant random slopes. Using a traditional ANOVA, we were able to replicate Gernsbacher's results for both younger and older groups with no effect of age group on suppression (see Table 3.6). Consistent with Gernsbacher's (1989) findings, both older and younger adults in this study tended to be slower in recognizing a probe after the anaphora when the probe was a nonreferent compared to a referent name. However, when considering the variances in the sensitivity to this interaction in each individual participant, the significant suppression and enhancement effects appeared to be eliminated. To make sure that the different findings of our and Gernsbacher's (1989) research was not due to the different participant inclusion criteria used, we ran a second analysis using only the data from participants that matched Gernsbacher's inclusion criteria (highly accurate participants). In this analysis, only younger participants were included. However, the results replicated our previous results and could not replicate Gernsbacher's (1989).

The findings from this study are consistent with the strategic rather than automatic theory of anaphora resolution (Greene et al., 1992), indicating that suppression of nonreferents and thus anaphora resolution can be strategic and postponed until required (e.g. when the comprehension question is asked). Whether participants postpone the suppression process might depend on the specific task and working memory demands. Since experimental sentences were short and working memory storage load was not very high in the task used in this study, it is possible that participants' working memory capacity was sufficient for maintaining all the information and concurrently storing and processing the new information in the discourse. Therefore, earlier suppression was not necessarily required in all trials,

which resulted in delayed resolution of anaphora. The effects of participants' performances in the preceding trials on their response times as well as accuracy also suggest that unrelated information was not completely inhibited by participants in this task. Although our results showed that older adults had a significant decline in comprehension accuracy, the underlying reason for such decline remains unclear. Further research is needed to determine whether the lack of suppression of nonreferents in the older adults in the current study was due to a strategic delay of resolution, as observed in the younger participants, or due to a cognitive deficit. A more demanding task that requires earlier suppression of information is needed to further investigate age-related differences in the process of anaphora resolution.

3.6 SUMMARY AND CONCLUSION

The study showed that older adults had lower accuracy in the comprehension of anaphoric pronouns relative to younger adults. Moreover, comprehension accuracy was affected by working memory capacity. Our findings suggested that working memory capacity contributed to individual differences in discourse comprehension abilities through affecting the ability to maintain the information that is part of the foundations for the mental representation of discourse. The advantage of first mentioned discourse characters in participants with higher working memory capacity demonstrated that they laid foundations for the mental representation. In contrast, participants with lower working memory capacity relied on recency and were not fully engaged in building the mental representation. Furthermore, we found no evidence for the suppression process during anaphora resolution either in the younger or in the older participants; by the time they finished reading the sentences. This suggested that anaphora resolution might be strategic rather than automatic and can be postponed under lower task demands.

Considering the difficulty faced by the older adults in comprehending potentially ambiguous pronouns, limiting the use of these linguistic devices in conversations involving older adults, and in written materials such as medication leaflets, and health and safety instructions, can help improve older adults' quality of life. Moreover, further investigation of anaphora resolution in older and younger adults using a more demanding task may lead to enhancements in our knowledge about language processing differences in older adults. A good understanding of how older individuals process language can help in employing ways to improve older adults' language comprehension.

4. CHAPTER FOUR

Age-related Changes in Anaphora Resolution under High Working Memory Storage Load

4.1 ABSTRACT

Chapter 3 reported a decline in accuracy in comprehending pronoun anaphora in older adults compared to younger adults. The analysis of the response times, however, revealed no effect of age group on the process of anaphora resolution. The current chapter investigated how ageing and working memory capacity affected anaphora resolution in a comprehension task with increased working memory load. As in Chapter 3, the influence of age and working memory on anaphora resolution was evaluated by examining advantage of first mention and suppression of irrelevant information in a probe recognition task. The same 30 younger and 30 older participants who performed Experiment 1 (Chapter 3) completed a second comprehension experiment on the same day. The experiment in this Chapter included a reading comprehension task with increased referential distance and a probe recognition task. Participants' comprehension accuracy and response times were analysed.

Analysis of data from this chapter showed that older adults were less accurate than younger adults in the comprehension of anaphoric pronouns. In addition, participants with higher working memory capacity performed better on the comprehension task. Analysis of the response times provided further evidence for the influence of working memory capacity on the advantage of first mention. Moreover, results demonstrated that, at the end of the sentence, younger adults' responses to referents were faster than responses to nonreferents. In older adults, however, there was no significant difference in response times to referents and nonreferents. The findings were interpreted to suggest that older adults did not suppress the accessibility of the nonreferents in resolving anaphoric pronouns. Taken together, the findings suggest that, although individual differences in working memory capacity contribute to differences in comprehension abilities, a decline in suppression abilities might mainly underlie age-related changes in comprehension difficulties.

4.2 INTRODUCTION

4.2.1 Current Study

The current study investigated how comprehension of anaphoric references was affected by ageing and working memory capacity under higher working memory storage load. In addition, it aimed to investigate whether the lack of suppression of the accessibility of nonreferents in older adults in Chapter 3 was due to delayed anaphora resolution or an age-related deficit in suppression abilities. Working memory storage load was increased through enlarging referential distance (two intervening sentences were added between referent and anaphoric pronoun). Under higher working memory storage load, more sophisticated processing is required as more unrelated information is to be suppressed. Specifically, we predicted that age-related deficit in suppressing the unrelated discourse information would underlie reduced performance in older adults compared to the younger adults.

The same paradigm as in Chapter 3 was used to address the following questions: 1) Do age and working memory capacity affect anaphora resolution in a comprehension task *under high working memory storage load*?; 2) Is comprehension accuracy more affected by increased working memory storage load in older compared to younger adults?; 3) Do age and working memory capacity affect advantage of first mention in a probe recognition task under a high working memory load?; and 4) Does age affect suppression of irrelevant information *under high working memory storage load?*

It was hypothesized that age and working memory would affect comprehension under high working memory storage load with: a) older adults being less accurate than younger adults in comprehension of anaphoric pronouns and b) higher working memory being associated with better comprehension. Moreover, increased working memory storage load was predicted to negatively affect comprehension accuracy in both older and younger adults.

However, this effect would be stronger in older adults resulting in a steeper decline if age-related comprehension difficulties were due to working memory capacity decline. It was further predicted that working memory capacity would affect the advantage of first mention under higher working memory storage load. First mentioned names would be expected to be more accessible than the second mentioned names in people with higher working memory capacity compared to those with lower working memory capacity. Given that older individuals might be expected to have lower working-memory capacity, age would also be expected to be related to the advantage of first mention. Finally, it is hypothesized that under high working memory storage load, younger adults would suppress the accessibility of nonreferents by the time they reach the end of the sentence. ageing is predicted to affect the suppression process. If older adults' comprehension decline is due to a decline in working memory capacity, they would be expected to suppress the nonreferents faster than the younger adults. In contrast, if age-related comprehension difficulties are due to inability to suppress the nonreferents, older adults would be expected not to suppress the nonreferents by the time they finish reading the sentences.

4.3 METHOD

The experiment described in this chapter (Experiment 2) was given to participants on the same day with Experiment 1 detailed in Chapter 3 and the Reading Span Task. The order of performing these experiments was counterbalanced across participants such that half of the participants performed Experiment 1 first and the other half performed Experiment 2 first.

4.3.1 Participants

The same participants completed Experiments 1 (Chapter 3) performed the current experiment (information regarding the participants was detailed in Chapter 3, Section 3.3.1).

4.3.2 Procedure

As stated, all individuals who participated in Experiment 1 (Chapter 3) completed two experiments (counterbalanced) at their visit to the University. Therefore, half of the participants performed Experiment 1 (See Chapter 3) first and the other half completed Experiment 2 first. As in Experiment 1, the comprehension experiment for this part was modelled on Gernsbacher (1989) and consisted of two subtasks: a reading comprehension task and a probe recognition task.

4.3.2.1 Experiment 2

In Experiment 2, participants were required to read a series of paragraphs on the computer monitor and answer the questions that followed. As they were reading the paragraphs, they were also periodically tested with a probe recognition task. Each comprehension paragraph consisted of four sentences; the first sentence introduced two characters one of which was referred to using an anaphoric pronoun in the last sentence (see Chapter 2, Section 2.3.2 for details of the experimental design). Details of the sentence stimuli are included below, followed by specifics of the experimental procedure. Information about the experimental probes and comprehension questions are detailed in Chapter 2, Sections 2.3.4 and 2.3.5.

4.3.2.2 Sentence Stimuli

The experimental stimuli were 72 English four-sentence paragraphs—48 were experimental sentences (See Appendix B), with the remaining 24 “lure” sentences. The 48 experimental sentence pairs were adopted from Gernsbacher’s (1989), experiment four. The current study only investigated the resolution of pronoun anaphora. Therefore, the noun anaphora of Gernsbacher’s sentences were substituted with pronoun anaphora. Removing the

effect of anaphora type from the study design provided us with a doubled number of trials in each condition. To increase the referential distance, two intervening simple sentences were added between the first and the last sentences in each sentence pair.

The first sentence in each paragraph contained two male or female characters, termed N1 (first person mentioned in the sentence) and N2 (second person mentioned in the sentence). One of the characters was referred to in the last sentence by an anaphoric pronoun (either *he* or *she*). The last sentence contained a subordinate clause followed by a main clause that contained the anaphoric pronoun. The subordinate clause was an adverbial participle clause that provided facilitating semantic information about the referent of the anaphoric pronoun. The two single-clause intervening sentences added between the first and the last sentences did not bias any of the names mentioned in the first sentences. Example sentences are as follows:

Example 1: Jodi (N1) picked up the washing for Beth (N2) before the first guests arrived. The house was very dirty. It has not been cleaned for almost a month. Glad to do the favour, she thought about the special friendship.

Example 2: Peter (N1) lost some money to Craig (N2) in gambling at the casino. The casino was located in a big hotel. It was always very busy. Enjoying the victory, he started walking toward the restaurant.

Stimulus sentences were constructed so that there were five words between N2 and the end of the first sentence, and five words between the anaphoric pronoun and the end of the last sentence. In half of the sentences, N1 was the reference of the anaphoric pronoun (as in Example 1) and in the other half N2 was the reference (as in Example 2). The second set of 24 sentences—all "lure"— were also adapted from Gernsbacher's (1989) Experiment 4. Names used in the probe task were not used in the stimuli sentences. Lure sentences used the

same structure as the 48 experimental sentences (12 with N1 as anaphora referent and 12 with N2 as anaphora referent).

4.3.2.3 Experimental Procedure for Experiment 2

As in Experiment 1, Experiment 2 also consisted of a comprehension task and a probe recognition task. Experimental procedures for the comprehension task are detailed in Chapter 2, Section 2.3.6. Participants were presented with a series of paragraphs followed by a comprehension question. While reading the paragraphs, participants were presented with a probe. General information about the probe recognition task is provided in Chapter 2, Section 2.3.7. Probes were displayed at two different testing points. Details of each testing point are as follows:

1. Before-Anaphora testing point: At the Before-Anaphora testing point, probes were presented 150 milliseconds after the offset of the last word of the last intervening sentence. Therefore, they were given right after the third sentence from the beginning of the paragraph (e.g., Bill poured John a cup of green tea. The green tea was of a very high quality. It was imported from Japan. *PROBE* Filling the cup too full, he spilled the tea all over.); or

2. After-Anaphora testing point: At the After-Anaphora testing point, probes were presented 150 milliseconds after the offset of the last word of the last sentence. In other words, they were given at the end of the paragraph (e.g., Bill poured John a cup of green tea. The green tea was of a very high quality. It was imported from Japan. Filling the cup too full, he spilled the tea all over.*PROBE*).

4.3.3 Data and Statistical Analysis

Two dependant variables were employed: (a) accuracy and (b) response time to the probe recognition task. For (a), two accuracy scores were calculated for each participant—both accuracy in response to the comprehension question and accuracy in completion of the probe recognition task. The data from 12 participants were removed as they completely or partially failed to perform one of the tasks. The removed data was the same data that was removed in the analysis of Experiment 1 (see Chapter 3, Section 3.3.3). In total, data from 27 out of 30 younger and 20 out of 30 older participants were included in the full analysis. Statistical information about participants' mean scores in comprehension and probe recognition tasks in Experiment 2 is provided in Table 4.1.

Table 4.1: *Mean accuracy in the comprehension and probe recognition tasks across participants (Experiment 2)*

(%) Accuracy	Group	Max	Min	Median	Mean	SD
Comprehension	Older	90.28	50.00	65.28	66.24	0.12
	Younger	98.61	72.22	88.89	86.93	0.07
Probe Recognition	Older	95.83	61.54	88.89	84.88	0.10
	Younger	100	80.56	95.83	94.65	0.04

For analysis of the RTs to the probe recognition task, only experimental trials that were responded to correctly (in both probe recognition and comprehension accuracy tasks) and with raw RTs greater than 300 milliseconds were included. The remaining data were then

further trimmed to exclude raw RTs larger than 2SD from the mean of all the correct experimental trials (1967 ms). These cut-offs removed 4.6% of the data. Removing the trials with incorrect comprehension accuracy excluded the possibility that response time patterns could be attributed to incorrect resolution of anaphora due to either information loss or switching the roles and the order of mention of the first and the second characters. Logarithmic transformation was used to normalize the RT distribution. Mixed effects modelling (MEM) was used to analyse comprehension accuracy and response times in Experiment 2. In addition, a separate binomial model was used to analyse the differences in the comprehension accuracy between the current experiment and Experiment 1 (Chapter 3).

4.4 RESULTS

4.4.1 Comprehension Accuracy

A series of binomial mixed effects models were used to analyze comprehenders' accuracy of responses to the comprehension questions. The effects examined included fixed effects and all the possible three-way and two-way interactions of the following factors: trial number, age group (younger or older), participants' genders, participants' years of education, gender of the sentence characters, participants' working memory capacity (Reading Span score), the position of the referent in the sentence (first or second), the testing point for the probe recognition task (before reading the anaphora or end of the sentence), log-transformed RT of the probe recognition task in the preceding trial, and the accuracy of response to both the preceding comprehension question and probe recognition task. All relevant interactions between these fixed effects were also tested. Random effects for participant and stimuli were also included. Including the random slopes in this model did not affect the significant results.

Model fitting was performed in a backward-stepwise fashion, followed by forward fitting of maximal random effects structure. Models were evaluated by model fitness

comparisons using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The final model for comprehension accuracy is presented in Table 4.2. Random effects included in the final model for comprehension accuracy are presented in Table 4.3.

Table 4.2: *Coefficients and p values of the binomial mixed effects model for comprehension accuracy in Experiment 2, with participant and stimulus as random effects.*

Effects	Coefficient	Standard Error	z value	P value
(Intercept)	-0.388	0.509	-0.763	0.445
Trial number	0.404	0.083	4.878	<0.001***
RP: Second	-0.595	0.196	-3.037	0.002**
Age group: younger	1.104	0.197	5.600	<0.001***
WM Capacity	0.030	0.009	3.361	<0.001***

*Note: WM = working memory, RP = referent position, * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$*

Table 4.3: *Random effects included in the final binomial mixed effects model for comprehension accuracy in Experiment 2.*

Random effects		Variance	Standard Deviation
Stimulus	(Intercept)	0.305	0.552
Participant	(Intercept)	0.209	0.457

The model revealed significant main effects of age group *($p < 0.001$)* and working memory capacity *($p < 0.001$)* on comprehension accuracy. Comprehension accuracy was

higher in the younger group compared to the older group. In addition, as working memory capacity decreased so did comprehension accuracy. The position of the referent in the sentence also significantly affected comprehension accuracy ($p < 0.01$). Participants were more accurate in finding the referents of the anaphora when the referent was the first mentioned name in the sentence compared to when it was the second mentioned name in the sentence. Furthermore, as expected, performance improved across the experiment, with a significant main effect of trial number ($p < 0.001$).

To analyse the effects of increased storage load in the second experiment compared to the first experiment on comprehension accuracy, a separate model was also fitted to the accuracy data from both Experiments 1 and 2. The effects examined included all previously examined effects and interactions plus the effect of experiment and all its possible interactions with the other effects. The final model and the random effects included are presented in Table 4.4 and Table 4.5 respectively.

Table 4.4: *Coefficients and p values of the binomial mixed effects model for comprehension accuracy including the effect of Experiment, with participant and stimulus as random effects.*

Effects	**Coefficient**	**Standard Error**	**z value**	**P value**
(Intercept)	-0.417	0.472	-0.884	0.377
Trial number	0.219	0.043	5.067	< 0.001***
RP: Second	-0.762	0.198	-3.857	< 0.001***
PTP: END	0.182	0.186	0.979	0.328
Age group: younger	0.772	0.190	4.066	< 0.001***
WM Capacity	0.028	0.008	3.466	< 0.001***

EXP: 1	-0.092	0.257	-0.360	0.719
Accuracy in PCT	0.170	0.129	1.316	0.188
EXP: 1*Accuracy in PCT	0.305	0.182	1.677	0.094·
RP: Second*EXP: 1	0.954	0.251	3.804	< 0.001***
RP: Second*Age group: younger	0.394	0.158	2.492	0.013*
PTP: END*EXP: 1	-0.540	0.251	-2.153	0.031*

*Note: EXP: Experiment, WM = working memory, PTP = probe testing point, PCT= preceding comprehension trial, RP: referent position, END = end of the sentence, · = p < 0.1, * = p < 0.05, ** = p < 0.01, *** = p < 0.001*

Table 4.5: *Random effects included in the final binomial mixed effects model for comprehension accuracy in both experiments.*

Random effects		**Variance**	**Standard Deviation**
Stimulus	(Intercept)	0.259	0.509
Participant	(Intercept)	0.203	0.450

The fixed effect of experiment was not significant. However, experiment had significant interactions with referent position ($p < 0.001$) and probe recognition testing point ($p < 0.05$). Participants were more accurate in finding the referents of the anaphora when the referent was the first mentioned name in Experiment 2 and when it was the second mentioned name in Experiment 1. Moreover, participants' accuracy was affected by referent position in Experiment 1 but not in Experiment 2. The interaction of experiment and the accuracy in the preceding trial was also approaching significance ($p < 0.1$) showing that accuracy was affected by the accuracy in the preceding trial in Experiment 1 but not in Experiment 2. Furthermore, the younger group and participants with higher working memory capacity were

generally more accurate ($p < 0.001$) and experiment had no interaction with age group or working memory capacity.

4.4.2 Response times

A series of linear mixed effects models were used to analyze participants' log-transformed response times to the probe recognition task. The effects examined in this analysis include the fixed effects of age group (younger or older), participants' genders, participants' years of education, gender of the sentence characters, participants' working memory capacity (Reading Span score), probe type (referent (REF) or nonreferent (NREF)), probe testing point (before anaphora or at the end of the sentence), referent position in the sentences (first or second), probe position in the sentences, trial number, log-transformed RT to the probe recognition task in the preceding trial, and the accuracy of responses to both the preceding comprehension question and probe recognition task. All relevant interactions between fixed effects were tested. Random effects for participant and stimulus were also included.

Model fitting was performed in a backward-stepwise iterative fashion, followed by forward fitting of maximal random effects structure. Model fitting was independently supported by model fitness comparisons using AIC and BIC. All significant results held when the fullest random effects structure was included. The final model for log-transformed RTs is presented in Table 4.6. This model included random intercepts for stimuli and participants, as well as by-participant random slopes for trial number, referent position, probe recognition testing point, probe type, and the interactions of probe type and referent position, and probe recognition testing point and probe type. Including these random effects allowed the model to take into account the mean differences in RTs across stimuli and across participants, as well

as variable sensitivity to the effects of task features across participants. The random intercepts and by-participant random slopes included in the final model are presented in Table 4.7.

Table 4.6: *Coefficients and p values of the final linear mixed effects model for RTs in Experiment 2, including fullest random effects structure.*

Effects	**Coefficient**	**Standard Error**	**t value**	**P value**
(Intercept)	7.271	0.162	44.900	< 0.001***
Trial Number	-0.053	0.010	-5.180	< 0.001***
Age Group: Younger	-0.278	0.062	-4.470	< 0.001***
RP: Second	0.088	0.103	0.850	0.395
PTP: END	0.023	0.030	0.790	0.430
WM Capacity	-0.002	0.003	-0.730	0.465
Probe Type: REF	0.067	0.091	0.740	0.459
Age Group: Younger*PTP:END	0.079	0.035	2.250	0.024*
Age Group: Younger*Probe Type: REF	-0.004	0.030	-0.130	0.897
RP: Second*WM Capacity	-0.002	0.002	-0.950	0.342
RP: Second*Probe Type: REF	-0.261	0.127	-2.050	0.040*
PTP: END*Probe Type: REF	-0.002	0.044	-0.040	0.968
WM Capacity*Probe Type: REF	-0.002	0.002	-1.020	0.308
Age Group: Younger*PTP: END*Probe Type: REF	-0.126	0.052	-2.440	0.015*
RP: Second*WM Capacity*Probe Type: REF	0.005	0.002	2.460	0.014*

*Note: WM = working memory, RP = referent position, PTP = probe testing point, REF = referent, END = end of the sentence, * = p < 0.05, ** = p < 0.01, *** = p < 0.001*

Table 4.7: *Random effects and slopes included in the final linear mixed effects model for RTs in Experiment 2.*

Random effects		**Variance**	**Standard Deviation**
Stimulus	(Intercept)	<0.001	0.018
Participant	(Intercept)	0.033	0.181
	Trial Number	0.003	0.051
	RP: Second	0.007	0.081
	PTP: END	0.006	0.075
	Probe Type: REF	0.002	0.044
	RP: Second*Probe Type: REF	0.007	0.079
	PTP: END*Probe Type: REF	0.019	0.135
Residual		0.031	0.176

Note: RP = referent position, PTP = probe testing point, REF = referent, END = end of the sentence

Question 2 in this chapter aimed to determine whether age and working memory capacity affect advantage of first mention in a probe recognition task with high working memory storage load. Relevant interactions testing these questions included that between working memory, referent type and referent position in the sentence, and that between age, referent type and referent position. The interaction between age, referent type and referent position was not significant and was removed from the model in the process of model development. Statistical analysis revealed a significant two-way interaction between probe type and referent position in the sentence *($p < 0.05$)*. Responses to referent probes were much faster when the referent was the first mentioned name in the sentence compared to when it was the second mentioned name in the sentence. Nonreferent probes were also recognized faster when the nonreferent was the first mentioned name in the sentence (probe type was

nonreferent and referent position was second) compared to when it was the second mentioned name in the sentence (probe type was nonreferent and referent position was first). This interaction is depicted in Figure 4.1.

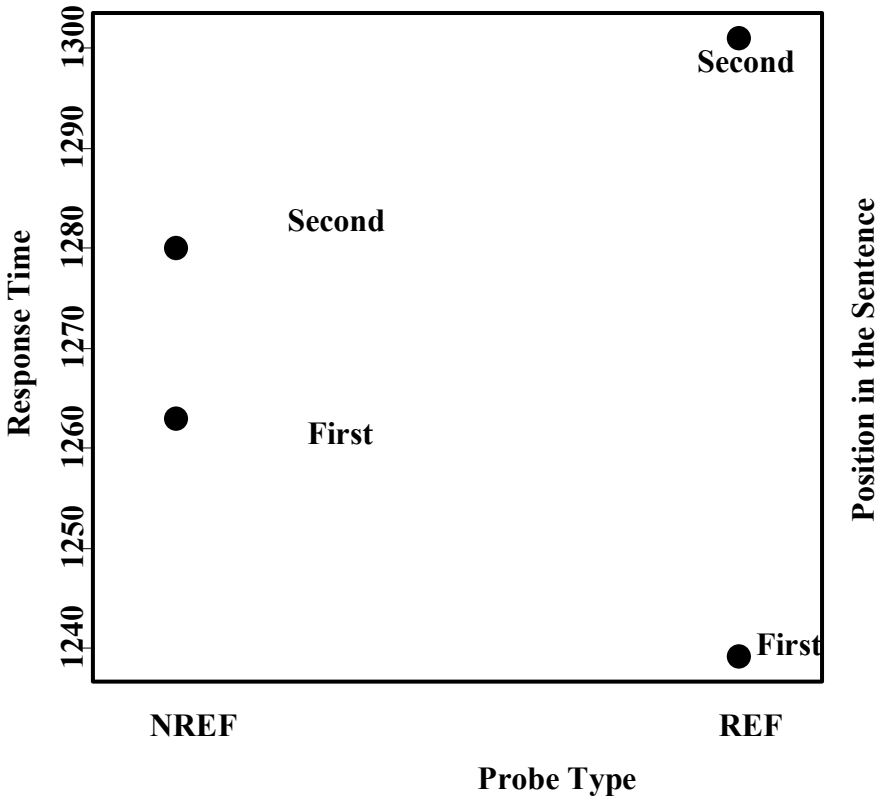

Figure 4.1: Two-way interaction of probe type and referent position

The interaction of probe type and referent position is further explained by the significant three-way interaction of working memory capacity, probe type and referent position *($p < 0.05$)*. Participants with higher working memory capacity were faster at responding to the probes when they were the first mentioned names in the sentence compared

to when they were the second mentioned names in the sentence. In other words, in participants with higher working memory capacity, responses were faster to the probes if they were referent and the referent was the first mentioned name in the sentence, or when they were nonreferent and the referent was the second mentioned name in the sentence. In contrast, participants with lower working memory scores had faster responses to the probes when they were the second mentioned names in the sentence compared to when they were the first mentioned names in the sentence. This interaction is depicted in Figure 4.2.

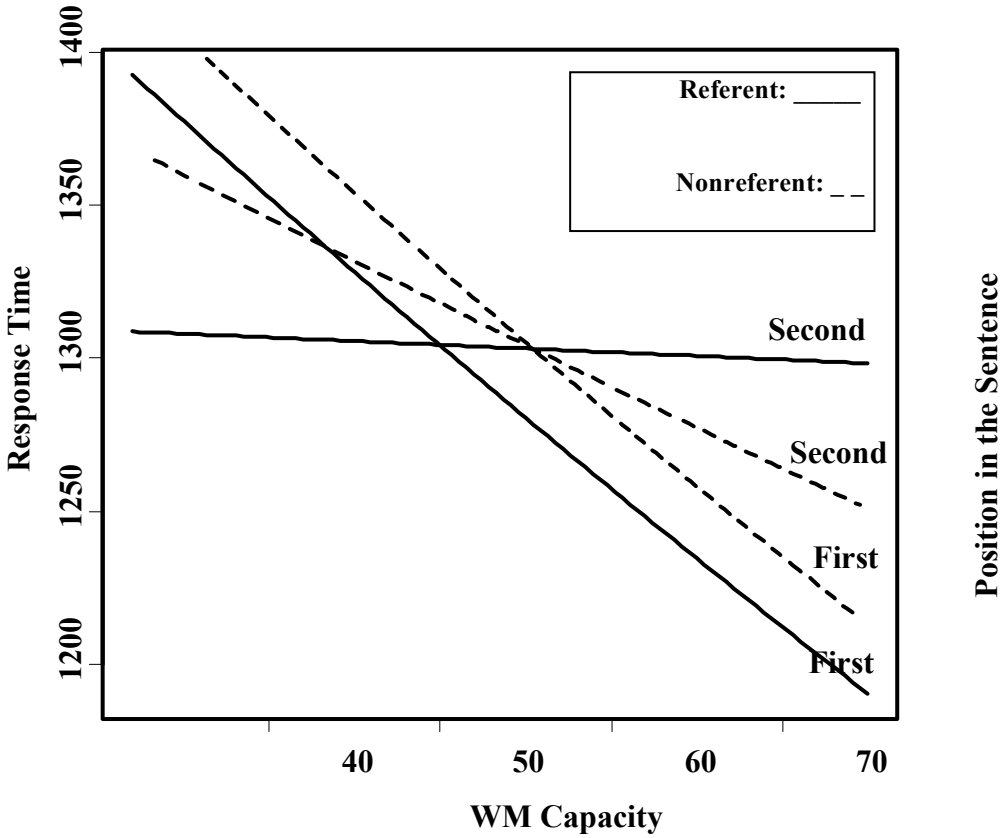

Figure 4.2: Three-way interaction of referent position, working memory capacity and probe type

Question 3 aimed to determine whether age affect suppression of nonreferent names during anaphora resolution under high working memory storage load. Relevant interaction testing this question was the interaction between age, probe type and probe testing point. statistical analysis revealed a significant interaction between age group and probe recognition testing point *($p < 0.05$)* showing that younger participants, but not older ones, were significantly slower in responding to probes at the end of the sentence compared to before anaphora. This interaction was further explained by the three-way interaction of age group, probe type and probe recognition testing point *($p < 0.05$)*. Younger participants' responses to the nonreferent probes were much slower compared to the referent probes at the end of the sentence. There was, however, no significant change in older adults' response times to referent and nonreferent names either before reading the anaphora or at the end of the sentence. This interaction in depicted in Figure 4.3.

Additional fixed effects included in the model that remained significant were age group *($p < 0.001$)* and trial number *($p < 0.001$)*. Older adults were significantly slower than younger adults in responding to the probes. Moreover, participants got faster in responding to the probes with the increasing number of trials.

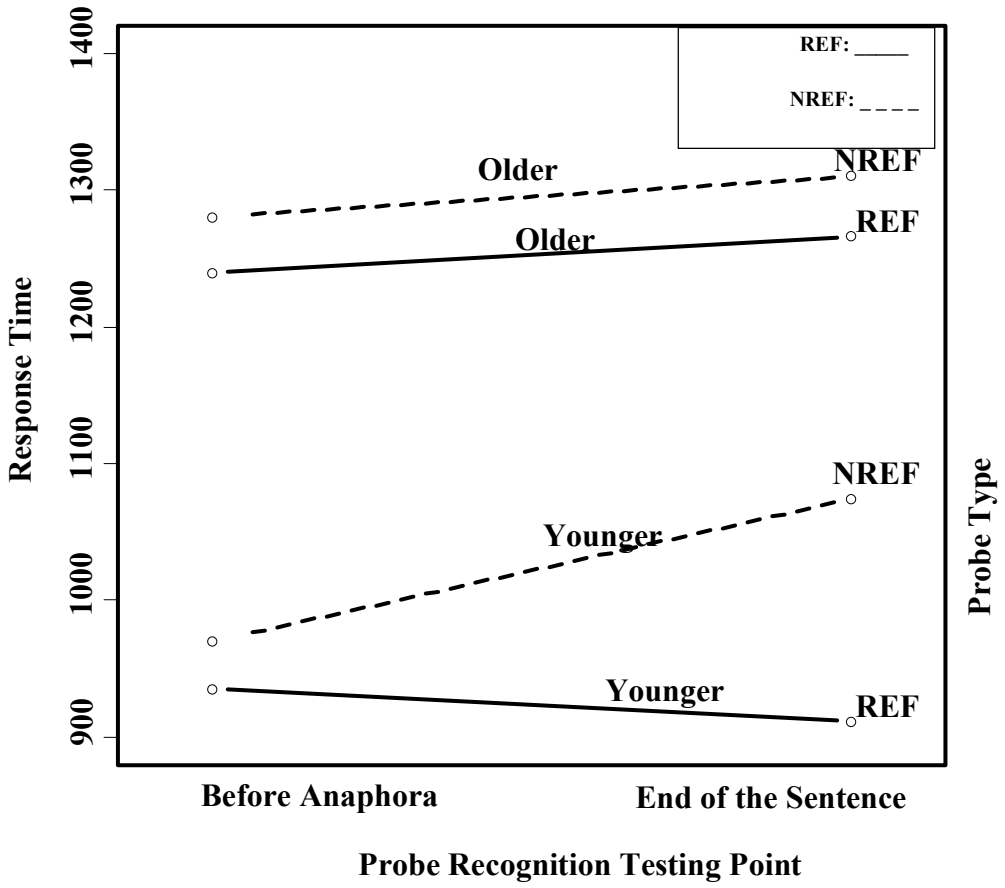

Figure 4.3: Three-way interaction of age group, probe type and probe recognition testing point

4.5 DISCUSSION

This study investigated the process of anaphora resolution under high working memory storage load and examined how age group and working memory capacity influenced this process. Consistent with the results from Chapter 3, the study found that age group and working memory capacity affected comprehension accuracy. Moreover, the results confirmed the earlier finding that only people with higher working memory capacity showed an advantage in response time to the first mentioned discourse characters. Unlike the results

from Chapter 3, the results from the current study showed a significant difference in the response times to referent and nonreferent names in younger adults at the end of the sentence. However, this interaction, which reflected the process of suppression, was absent in the older adults. These particular results are discussed in relation to the findings from Chapter 3 in separate sections.

4.5.1 Effects of Age and Working Memory Capacity on Comprehension Accuracy

The results from the current experiment replicated two main findings from the analysis of comprehension accuracy scores in Experiment 1 (see Chapter 3, Section 3.5.1). First, the results showed that age negatively affected anaphora resolution accuracy. Second, the results confirmed that higher working memory scores were associated with improved language comprehension in both older and younger groups. It was suggested in Chapter 3 that working memory capacity and ageing might contribute separately to individual differences in the underlying processes involved in anaphora resolution. That is the effect of ageing on language comprehension might not merely be due to the differences in working memory capacity. Rather, other cognitive deficits might have contributed to the age-related changes in anaphora resolution. Response time data supported this hypothesis showing that working memory capacity, but not ageing, affected the advantage of first mention in the process of anaphora resolution. This hypothesis was further investigated in the current experiment by analyzing the response time data (see Sections 4.5.2 and 4.5.3)

Moreover, it had been hypothesized that if age-related changes in anaphora resolution were due to working memory capacity decline, increasing the storage load would result in a steeper accuracy decline in older adults compared to the younger adults. Capacity-based theory of cognitive ageing suggested that in older adults more capacity was devoted to processing and thus storage was deficient resulting in the loss of information in older adults.

Therefore, it was expected that increasing the amount of information affected older adults' accuracy more than younger adults'. In this case, older adults were expected to show a steeper decline compared to younger adults in comprehension accuracy in Experiment 2 compared to in Experiment 1. However, the analysis of comprehension accuracy including data from both experiments did not reveal an interaction between age and experiment. This result suggested that increased working memory storage load equally affected comprehension accuracy in younger and older adults.

In contrast to the findings from Experiment 1 (low storage condition from Chapter 3), the results of the current experiment indicated that comprehension accuracy was not influenced by other task features such as probe recognition testing point and accuracy in the preceding trials. This finding suggests that the higher working memory storage load in this task might have required the participants to focus more attention on the specific task goals and thus participants were less distracted by goal-irrelevant factors. In general, the comprehension accuracy results from Experiments 1 and 2 showed that, regardless of task complexity and working memory demands, comprehension accuracy was affected by age and working memory capacity. The effects of these factors on the process of anaphora resolution were further investigated through the analysis of response times to the probes in the probe recognition task. The analysis of RTs revealed different contributions of age and working memory capacity to the changes in the process of anaphora resolution. Results indicated that working memory capacity affected the advantage of first mention, which was an index of the process of laying the foundation while ageing affected suppression in the processes of information update during anaphora resolution.

4.5.2 Effects of Age and Working Memory Capacity on Advantage of First Mention

Firstly, the results from the current experiment provided further evidence for the discourse phenomenon of the advantage of first mention during discourse processing. That is, participants were more accurate in inferring the referent of the pronoun anaphora when it was the first mentioned name in the sentence. This finding is consistent with the results from the analysis of response times in this experiment as well as with the findings from the earlier studies showing an advantage for the first mentioned discourse entities (Gernsbacher, 1989; Gernsbacher & Hargreaves, 1988; Järvikivi et al., 2005; McDonald & Shaibe, 2002).

The interaction between referent position and probe type found in the analysis of RTs in this study showed that first mentioned names were recognized faster than the second mentioned names. In other words, in the probe recognition task, responses were significantly faster when probes were referent names and the referent was in the first position in the sentence or when probes were nonreferent names and the referent was in the second position. The responses were slower when probes were referent names and the referent was in the second position in the sentence, or when probes were nonreferent names and the referent was in the first position in the sentence. Therefore, first mentioned names were more accessible than the second mentioned names.

Secondly, the results further supported the finding that working memory capacity affected the advantage of fist mention. Participants with higher working memory capacity were faster at responding to the probes when the probes were the first mentioned name in the sentence compared to when they were the second mentioned name in the sentence. In contrast, participants with lower working memory scores responded faster to the probes when they were the most recent names in the sentence (second mentioned name). Several studies have shown that the advantage of first mention is because the first mentioned discourse entities form the foundation of the mental representation for comprehension (Carreiras et al.,

1995; Gernsbacher & Hargreaves, 1988; Kim et al., 2004). Since this advantage has been suggested to demonstrate the process of laying the foundation for mental representation of discourse, working memory capacity was proposed to affect the process of laying the foundation (See Chapter 3, Section 3.5.2).

Consistent with the findings from Chapter 3 and the study of van Rij and colleagues (2013), the results from the current experiment further demonstrated that participants with higher working memory capacity were better able to lay foundations for comprehension. In contrast, those with lower working memory capacity relied more on the recency of the discourse concept. Moreover, as in the first experiment (see Chapter 3), age group did not affect the advantage of first mention. However, an ageing effect was found on the process of suppression. Significant effect of age, but not working memory capacity, on suppression process supported the hypothesis that age and working memory capacity might contribute independently to the individual differences in anaphora resolution.

4.5.3 Effects of Age on Suppression Process

An important aim of the current study was to investigate the effects of age on the process of anaphora resolution. Based on the inhibitory decline theory of cognitive ageing, it was particularly hypothesized that inefficient suppression of nonreferents in the discourse underlies the age-related changes in the ability to comprehend anaphora. The suppression process was investigated through measuring the change of accessibility of referent and nonreferent names from before anaphora to the end of the sentence.

In contrast to the results obtained from the first experiment, the analysis of response times in the current experiment demonstrated suppression of nonreferent names by younger adults by the time they finished reading the sentences. At the end of the sentence, younger adults were faster in recognizing the referent names than the nonreferent names. This finding

was consistent with Gernsbacher's (1989) model of anaphora resolution suggesting that pronoun anaphora resolution is facilitated by suppressing the nonreferents while enhancing the referent. This finding also supported the hypothesis that working memory load affected the time-course for anaphora resolution. While younger adults delayed the process of anaphora resolution under low working memory load in Experiment 1, they completed the process earlier under higher working memory load in Experiment 2 so that the irrelevant information could be suppressed earlier. Earlier suppression was suggested to provide more capacity for storing and processing the remaining of the sentence.

It was argued that if age-related comprehension difficulties were only due to working memory capacity decline, older adults would be expected to suppress the nonreferents earlier than the younger adults did. Older adults' response times, however, did not reveal any significant difference between the response times to referent and nonreferent names either before anaphora or at the end of the sentence. This finding suggested that the older adults did not suppress nonreferent names by the time they finished reading the sentences. Successful anaphora resolution requires the referent to become more accessible than the nonreferents in the comprehender's mind through suppression and enhancement processes. The lack of suppression in the process of anaphora resolution in the older adults could account for their lower comprehension accuracy compared to the younger adults.

In the current analysis, only trials with correct answers to the comprehension questions were included which excludes the possibility that participants forgot the names of the characters or switched their roles. Therefore, these factors could not be the reason underlying the lack of suppression and the inability to recognize a single referent. Moreover excluding the data from participants with low comprehension accuracies reduced the possibility that answers were based on chance. Three main reasons for the lack of suppression

in older adults could be put forwards: 1) an inhibitory decline, 2) general slowing, and 3) use of a different strategy.

It was hypothesized that older adults' poorer comprehension was due to an inability to suppress the unrelated information in the discourse. The lack of suppression in the current study was consistent with this hypothesis and could provide partial support for the inhibitory decline theory of cognitive ageing proposed by Hasher and Zacks (1988). This theory suggested that an age-related inhibitory deficit might underlie the changes in language processing in ageing. However, although this theory considers a decline in all inhibitory abilities to underlie cognitive ageing, results from the current study showed that only suppression abilities were consistent with the inhibitory decline theory of cognitive ageing.

Based on the general slowing theory of cognitive ageing (Salthouse, 1996), older adults take longer to process information compared to younger adults. When the processing time is not self-controlled by older adults, increasing information load would result in simultaneous processing of the previous and new information. This would also require longer maintenance of information, which along with simultaneous processing would impose high demands on older adults' limited working memory capacity. Exceeding working memory capacity results in the loss of information and therefore impaired comprehension. The results of the current investigation appear to provide some support for this theory—it might be argued that lack of suppression at the end of the sentence was due to slowed processing. However, in this case we would further expect the high working memory storage load in Experiment 2 to affect older adults' comprehension accuracy more than younger adults. In other words, a comparison between the comprehension accuracy in Experiment 2 and in Experiment 1 would reveal a steeper decline in older adults' accuracy compared to the younger adults'. However, the results from the analysis of comprehension accuracy in the

two experiments suggested that both younger and older adults were equally affected by increased working memory storage load.

Another alternative interpretation of the results might attribute the age effect on suppression to the use of different strategies by younger and older adults. In Chapter 3, it was suggested that anaphora resolution could be strategic and delayed. In the current experiment, we expected to see a change of strategy (immediate rather than delayed resolution) due to high working memory demands. It might, therefore, be argued that the difference in the patterns of accessibility in older and younger adults reflected applications of different strategies for anaphora resolution. In other words, younger adults changed their strategy from first to second experiment to keep their high level of accuracy under higher working memory capacity. In contrast, older adults did not change their strategy for the second experiment and delayed the resolution in both Experiments 1 and 2. However, as in general slowing hypothesis, in this case we would expect a significant interaction between age group and experiment showing a steeper decline in accuracy in the second experiment compared to the first experiment in older adults compared to younger adults. However, as mentioned earlier, this interaction was not significant. Therefore, it seems that older adults' lack of suppression cannot be attributed to a different strategy. Considered within the context of the current data set, it would appear that inhibitory decline theory proposed by Hasher and Zacks (1988) provides the most feasible interpretation, given that it could best explain the age-related changes in both processing of anaphora and comprehension accuracy. However, current results suggest that the inhibitory decline theory might be restricted to suppression abilities.

Overall, the findings from this study highlighted the important roles for both working memory capacity and suppression in anaphora resolution. The current results suggested that individual differences in both working memory capacity and suppression abilities might

contribute to the individual differences in discourse processing. Although working memory capacity measured with reading span task affected comprehension scores, it might not reflect the age-related differences in anaphora resolution. Age-related differences were suggested to be due to deficient suppression of irrelevant discourse information.

It might be argued that working memory capacity, as an index of the ability to maintain information in the presence of distractions, should predict any differences in inhibitory abilities. That is because inhibitory functions play crucial roles in making an efficient use of the limited capacity, particularly under distracting situations. Therefore, individual differences in working memory capacity might also be expected to contribute to the ability to suppress the irrelevant discourse information. However, this effect was not significant in the current study. This could be explained by the differences between the inhibitory abilities involved in performing the reading span task and those that are crucial to discourse processing. Reading span task is a measure of the ability to control attention to maintain information in a distracting context. In addition to these abilities, discourse processing requires the ability to maintain the information that is task-relevant by ongoing update of the active information (Was et al., 2011). Although individual differences in maintaining information in the presence of task-irrelevant distractions affected language comprehension, age-related changes in language comprehension were mainly due to differences in the ability to suppress the task-relevant information. This finding supports the existence of different inhibitory functions proposed in previous studies (Blasi et al., 2006; Brydges et al., 2012; Friedman & Miyake, 2004b; Gernsbacher, 1997b; Lustig et al., 2007) and that inhibitory functions might be differently affected by psychological and neurological diseases and normal ageing (Friedman & Miyake, 2004b).

Our findings suggest a distinction between inhibitory functions that control the activation of distracting information and those that suppress the previously activated information that is no longer relevant to the goals of the task. Current study results are consistent with Bell and colleagues' (2008) study. These authors suggested that age-related comprehension difficulties were restricted to inhibitory abilities responsible for suppression of the information which used to be relevant to the goals of the comprehension task before further information made them goal-irrelevant (Bell et al., 2008). Comparing the results from Experiment 1 and Experiment 2 in the current study, it seems that both older and younger adults were similarly affected by goal-irrelevant distractions. In Experiment 1, both comprehension accuracy and response times were affected by the performance in the preceding trial and the testing points which were irrelevant to the goal of the comprehension task (i.e. answering the comprehension question). In Experiment 2, however, the performance was not affected by these factors in either group, which showed stronger inhibition of distractions because of higher working memory storage load. This suggested that older adults might have been as capable as younger adults in inhibiting goal-irrelevant distraction. However, they were not able to suppress the nonreferents, which were goal-relevant and only became irrelevant when further information was received.

4.6 SUMMARY AND CONCLUSION

This study provided further evidence that ageing might negatively affect accuracy of anaphora resolution. Moreover, working memory capacity had positive influence on comprehension accuracy. Increased working memory load in this experiment compared to Chapter 3 similarly affected younger and older adults and resulted in poorer performance on comprehension task regardless of age. Results from recognition task provided further support for the effect of working memory capacity on the advantage of first mention. In addition,

results demonstrated suppression of nonreferents in younger adult. However, suppression was not observed in older adults. Findings suggest that age-related changes in anaphora comprehension might be mediated by a decline in inhibitory abilities responsible for suppression of the previously relevant information, which became irrelevant to the task goals upon receiving further information. Overall, findings demonstrated that individual differences in working memory capacity and ageing affected different sub-processes involved in anaphora resolution. While ageing affected the suppression process, working memory capacity influenced the advantage of first mention, which had been suggested to demonstrate the process of laying a foundation for mental representation of discourse. Moreover, it was suggested that time-course for suppression might depend on task's working memory demands.

4.7 FURTHER DIRECTIONS

Taken together, the results from Chapters 3 and 4 demonstrated that anaphora resolution might be strategic and dependant on task features. While younger participants resolved anaphora by the time they reached the end of the sentence in Experiment 2 (current chapter), they delayed this process in Experiment 1 (see Chapter 3). It was suggested that earlier anaphora resolution in Experiment 2 indicated a need for earlier suppression of the irrelevant information to allow larger capacity for storage and processing of further information. Increased amount of information in Experiment 2 might have exceeded the limits of working memory capacity, requiring older information to be suppressed from the limited storage capacity. If working memory capacity was not used efficiently through suppressing the irrelevant information, increasing the amount of information could result in either the loss of important older information or inability to store and process the new information. Therefore, anaphora might have been resolved earlier so that the unrelated

information could be determined and suppressed. This suggests that besides the crucial role of resolving ambiguity, suppression process might play important roles in unambiguous discourse comprehension under demanding conditions. If earlier suppression in this study was due to higher storage load, it could be argued that suppression ability contributed to better discourse comprehension through improving storage efficiency.

However, it is possible that factors other than increased storage load contributed to the earlier resolution of anaphora in the Experiment 2. Two alternative factors might have contributed to earlier suppression of nonreferents; higher processing load, and prior disambiguation. Investigating the main reason for earlier anaphora resolution in the second experiment provides an insight into how suppression of irrelevant discourse information could contribute to keeping comprehension accurate. Moreover, it might help understanding how an inability to suppress the irrelevant discourse information in ageing could contribute to age-related decline in discourse comprehension.

First, the sentences in the second experiment were syntactically different from those used in the first experiment. Sentences containing anaphoric pronouns were simple sentences in Experiment 1 and complex sentences in Experiment 2. Differences in syntactical complexity of sentences involved in a comprehension task are suggested to contribute to differences in processing demands of the comprehension tasks (Prat, Keller, & Just, 2007). There is some evidence that syntactically complex sentences are more difficult to process than simple sentences (Prat et al., 2007).Therefore, besides storage load, processing load might also be argued to be higher in the second experiment. Thus, processing the complex sentences in Experiment 2 might have imposed higher processing demands on participants' working memories. This hypothesis is consistent with the capacity-based theory of working memory. It might be argued that to devote larger capacity to processing under high

processing load, smaller capacity should be devoted to storage. Therefore, older irrelevant information needed to be suppressed earlier resulting in earlier anaphora resolution. If the irrelevant information was not suppressed from working memory, smaller capacity would have remained for processing which might have not been sufficient for processing of complex sentences. This would suggest that suppression ability contributes to higher comprehension accuracy through improving information processing abilities.

Second, in the experimental sentences used in the second experiment, the disambiguating information necessary for anaphora resolution was provided before anaphora (e.g. *Peter lost some money to Craig in gambling at the casino. The casino was located in a big hotel. It was always very busy. Enjoying the victory, he started walking toward the restaurant.*), while this information was provided after anaphora in the first experiment (e.g., *Angela gave Nicole some directions to the zoo and she had no trouble following them.*). The disambiguating contextual information combined with the information provided by the anaphoric pronoun helps in the resolution of anaphora and triggers suppression (Gernsbacher, 1989). Prior disambiguation in the second experiment might have been facilitating and thus helped in earlier resolution of anaphora. This could suggest that prior disambiguation rather than earlier suppression contributed to better comprehension.

In summary, since ageing mainly affected the suppression process, it seems crucial to further investigate the separate contributions of the increased processing load, high storage load, and prior disambiguation to the earlier suppression of nonreferents as well as to comprehension accuracy. If earlier suppression in younger adults is due to a facilitating effect of prior disambiguation, comprehension accuracy will be expected to be higher when contextual information is provided earlier. In such case, older adults' lack of suppression could be due to an inability to make use of this facilitating information. In contrast, if earlier

suppression in younger adults is caused by higher processing or storage demands of the task, comprehension accuracy might be expected to decline with increased storage or processing load. In such case, earlier suppression would be considered a compensating strategy to keep comprehension accurate by either improving the storage or processing abilities. Older adults' lack of suppression would also be expected to result in either storage or processing deficits. Therefore, investigating the main cause for younger adults' earlier suppression in Experiment 2 (see Chapter 4) can help in better understanding the underlying cognitive mechanisms of discourse comprehension and the age-related changes in this process.

5. CHAPTER FIVE

A Follow-up Investigation into the Mechanisms Underlying Anaphora Resolution

5.1 ABSTRACT

Results from Chapter 3 suggested that under lower memory storage load, younger adults postponed the resolution of anaphora—as evidenced by the lack of suppression of the accessibility of nonreferent name when accessibility was measured at the end of the sentence. Under higher storage load in Chapter 4, however, younger participants suppressed the accessibility of nonreferent name by the time they reached the end of the sentence. While this could be interpreted to mean that higher storage load required irrelevant information to be suppressed earlier from working memory, it is possible that other factors contributed to the earlier suppression. These factors include higher processing demands of the task and the facilitating effect of prior disambiguation of pronoun anaphora. Given this possibility, the current study aimed to determine the contributions of task's storage demands, task's processing demands, and earlier disambiguation to earlier suppression of the accessibility of nonreferents during anaphora resolution.

Forty younger participants completed four comprehension experiments and a reading span task. The independent variables of working memory storage load (as manipulated by increasing referential distance), processing load (as manipulated by using more complex sentences) and time-course (prior versus after anaphoric pronoun) for providing contextual information were manipulated across these experiments. Each of the four comprehension experiments included a reading comprehension task and probe recognition task. Sentences used in the reading comprehension task included an anaphoric pronoun, a referent and a nonreferent name. Similar to Chapters 3 and 4, the dependent variables included participants' comprehension accuracy and response times in the recognition task. Response times were used to measure how accessible the referent and the nonreferent names were in working memory, and were collected either before participants encounter anaphora or at the end of the

sentence. Results of these studies found that task's processing demands negatively affected comprehension accuracy and resulted in earlier suppression of nonreferents. In contrast, suppression process was not observed in tasks with higher storage demands. Nor did storage load affect comprehension accuracy. Moreover, accuracy was lower in sentences in which disambiguating information was provided prior to anaphoric pronoun, ruling out the facilitating effect. It was suggested that higher processing load rather than higher storage load and earlier disambiguation contributed to earlier suppression of nonreferents during anaphora resolution. The current study findings suggest that suppression of irrelevant discourse information might contribute to better discourse comprehension under high processing load through improving processing abilities.

5.2 INTRODUCTION

Chapters 3 and 4 reported a reduced accuracy in the comprehension of pronoun anaphora in older adults compared to the younger adults. Analysis of response time data revealed that younger adults suppressed the nonreferent name during anaphora resolution under high working memory storage load. In contrast, older adults were unable to suppress the nonreferent name in the process of anaphora resolution under both high and low storage loads. The ageing effect on the process of anaphora resolution was only found in the second experiment, which had increased referential distance and thus higher working memory storage demands.

Younger adults resolved anaphora by the time they finished reading the sentences under high working memory storage load (Experiment 2) but delayed anaphora resolution under lower working memory storage load (Experiment 1). It was suggested that when working memory demand was lower, anaphora resolution might be delayed, as readers' working memory capacity was sufficient to maintain all the information for later processing.

In such case, earlier suppression of information was not necessary and anaphora could be resolved when required by the comprehension question. It was also suggested that higher working memory storage load in the second experiment might have exceeded working memory capacity limits and thus readers were required to suppress the unrelated information, which resulted in earlier resolution of anaphora. An inability to suppress the irrelevant information could result in the loss of important information or deficient storing of the new information. It might thus be argued that earlier anaphora resolution in this case contributed to an efficient use of working memory limited storage capacity and improved storing abilities.

However, two alternative explanations are also possible for the earlier suppression in Experiment 2 (higher working memory load, see Chapter 4) compared to in Experiment 1 (lower working memory load, see Chapter 3); higher processing demands, and prior disambiguation. The anaphoric sentences used as experimental stimuli in these two experiments were different in terms of syntactic complexity and time-course for providing disambiguating contextual information. Providing the contextual information through an adverbial participle clause before anaphoric pronoun might have accounted for the faster anaphora resolution in Experiment 2 (e.g. *Peter lost some money to Craig in gambling at the casino. The casino was located in a big hotel. It was always very busy. Enjoying the victory, he started walking toward the restaurant*.). On one hand, it is possible that earlier disambiguation facilitated anaphora resolution and resulted in earlier update in working memory and suppression of the nonreferent name. In this case, facilitation would be expected to result in increased comprehension accuracy. On the other hand, earlier suppression of the nonreferent name observed in Experiment 2 (higher working memory load, see Chapter 4) might have been due to higher processing demands for comprehending syntactically complex sentences

In Experiment 1 (lower working memory load, see Chapter 3), anaphora was included in a simple clause that was part of a compound sentence. In Experiment 2 (higher working memory load, see Chapter 4), however, anaphora was included in a main clause, which was part of a complex sentence starting with a subordinate clause. There is some evidence that complex sentences are more demanding and more difficult to process than compound sentences (Prat et al., 2007). It is, therefore, possible that higher processing demands in Experiment 2 contributed to earlier suppression. In such case, higher processing demands would be expected to result in decreased comprehension accuracy. Earlier suppression of irrelevant information would be required so that larger capacity could be available for processing of information and thus improve processing abilities. This hypothesis is consistent with shared-resources approaches to working memory capacity. Further investigation of the factors contributing to younger adults' earlier suppression in Experiment 2 can shed light on how suppression of irrelevant information might contribute to keeping comprehension accurate in more demanding tasks.

5.2.1 Current Study

This part of the current study aimed to follow up on the findings from Chapters 3 and 4. Two more experiments (Experiments 5 and 6) were designed to determine whether earlier resolution of anaphora was due to prior disambiguation, higher processing demands of the more syntactically complex sentences, or higher working memory storage demands. Participants were required to complete the new experiments in addition to the short versions of the experiments used in Chapters 3 and 4 (Experiments 3 and 4).

To separately investigate the effects of working memory storage load and syntactic structure of the anaphoric sentences on comprehension accuracy and suppression process, these factors were manipulated in Experiments 5 and 6. These new experiments included: 1)

a comprehension task with low storage load in which contextual information was provided before the anaphoric pronoun, and 2) a comprehension task with high working memory storage load in which contextual information was given after anaphoric pronoun. The sentences containing anaphora were simple sentences in the Experiments in which contextual information was provided after the anaphora. The anaphoric sentences were complex in the experiments in which contextual information was provided before the anaphora. Information about working memory storage demands, syntactical complexity of anaphoric sentences and time-course for disambiguation in each experiment is presented in Table 5.1.

Table 5.1: *Features of the experiments used in chapter 5*

	Low storage Load	**High storage Load**
Early disambiguation **More Syntactically Complex**	Experiment 6	Experiment 4
Late disambiguation **Less Syntactically Complex**	Experiment 3	Experiment 5

If earlier anaphora resolution in Experiment 2 was due to higher working memory storage load, we would expect to observe suppression of nonreferents under a high storage load condition in a less syntactically complex task with later disambiguation (Experiment 5). In contrast, if the nonreferent name is suppressed under a low storage load in more syntactically complex task with earlier disambiguation (Experiment 6), it is possible that earlier disambiguation or higher processing demands contributed to earlier suppression. In the latter case, if the suppression is a result of facilitation caused by earlier disambiguation, we

would expect higher accuracy in the experiments in which contextual information is provided earlier. However, if the suppression is a result of higher processing demands, we would predict a decline in comprehension accuracy in these experiments.

The following questions were addressed in the current study: 1) How is comprehension accuracy affected by working memory storage load and the syntactic structure of the sentences?; 2) Do younger adults suppress the accessibility of the nonreferents by the time they reach the end of the sentence, in simpler sentences with increased storage load and late disambiguation?; and, 3) Do younger adults suppress the accessibility of nonreferents by the time they reach the end of the sentence, in more syntactically complex sentences with low storage load and prior disambiguation?

It was hypothesized that: 1) if earlier suppression was due to higher processing demands, comprehension accuracy would be lower in more syntactically complex sentences with prior disambiguation. In contrast, if earlier suppression was due to facilitating effect of prior disambiguation, comprehension accuracy would be higher in these sentences; 2) if earlier suppression was due to high storage load, in simpler sentences with high storage load and late disambiguation, the accessibility of nonreferents is suppressed when measured at the end of the sentence.; and, 3) if earlier suppression was due to different syntactic structures, in more syntactically complex sentences with low storage load and prior disambiguation the accessibility of the nonreferents is suppressed when measured at the end of the sentence.

5.3 METHOD

The current study includes two new comprehension experiments. These two experiments were given to the participants along with the shorter versions of the comprehension experiments used in Chapters 3 and 4.

5.3.1 Participants

Participants included 40 (15 males and 25 females, *M = 22.97, SD = 4.96*, range = 18-38 years) right-handed native speakers of New Zealand English (NZE). They reported no history of neurological disease, dementia, cognitive impairment, cardiovascular disease, uncontrolled hypertension, learning disability, attention deficit disorder, or speech disorder.

5.3.2 Procedure

Participants in this study completed five tasks: four comprehension experiments (Experiments 3, 4, 5, and 6) followed by an assessment of working memory capacity—the Automated Reading Span Task (Unsworth et al., 2005). The comprehension experiments were modelled on Gernsbacher (1989) and consisted of two subtasks: a reading comprehension task and a probe recognition task. Participants performed all the experiments on the same day, with a break after each experiment. All participants performed Experiments 3 and 4 before Experiments 5 and 6. Half of the participants performed Experiment 3 first and the other half performed Experiment 4 first. Moreover, half of the participants completed Experiment 5 before the Experiment 6 and the other half completed them in a reverse order.

5.3.2.1 Comprehension Experiments

Participants performed four comprehension experiments, each of which took between 20 to 25 minutes. In all experiments, they were required to read a series sentences/paragraphs on the computer monitor and answer the questions that followed. As they were reading the sentences/paragraphs, they were also periodically tested with a probe recognition task (see Chapter 2, Section 2.3.2 for details of the experimental design). Details of the sentence stimuli in each experiment are included below, followed by specifics of the experimental

procedures. Information about the experimental probes and comprehension questions are detailed in Chapter 2, Sections 2.3.4 and 2.3.5 respectively.

5.3.2.2 Sentence Stimuli

a) Experiment 3

The experimental stimuli were 50 English sentences—32 were experimental sentences, with the remaining 18 lure sentences. Experimental sentences were chosen from the experimental sentences used in Experiment 1 (See Chapter 3, Section 3.3.2.2). In half of the experimental and lure sentences, N1 was the reference of the anaphoric pronoun and in the other half N2 was the referent.

b) Experiment 4

The experimental stimuli were 36 English four-sentence paragraphs—24 were experimental sentences, with the remaining 18 lure paragraphs. Experimental paragraphs were chosen from the experimental paragraphs used in Experiment 2 (See Chapter 4, Section 4.3.2.2). In half of the experimental and lure paragraphs, N1 was the reference of the anaphoric pronoun and in the other half N2 was the referent.

c) Experiment 5

The experimental stimuli were 36 English four-sentence paragraphs—24 were experimental paragraphs (See Appendix C), with the remaining 12 lure paragraphs. 24 experimental sentence pairs were adopted from Gernsbacher's (1989) two-clause sentences. Each two-clause sentence was broken into two separate simple sentences to form a sentence pair. To increase the referential distance, two intervening simple sentences were also added between the two sentences in each sentence pair (see Example 1 below).

Example 1: Pamela picked Stacey some yellow and white flowers. The flowers were all roses. They smelt very nice. She gathered a very big bouquet.

The first sentence in each paragraph contained two male or female characters, termed N1 (first person mentioned in the sentence) and N2 (second person mentioned in the sentence). One of the characters was referred to in the last sentence by an anaphoric pronoun (either *he* or *she*). The last sentence was a simple sentence starting with the anaphoric pronoun. The remaining of the last sentences provided contextual information required to resolve anaphora. The two single-clause intervening sentences added between the first and the last sentences did not bias any of the names mentioned in the first sentences. Example sentences are as follows:

Example 1: Jennifer (N1) made Michelle (N2) a Chinese dish for lunch. The dish was very popular in China. It was mainly made of fish and rice. She used an old fashioned recipe.

Example 2: Darrell (N1) told Gabriel (N2) about the new short movie. It was made by a very famous director. Great actors had played in the movie. He enjoyed hearing the movie review.

Stimulus sentences were constructed so that there were always five words between N2 and the end of the first sentence, and five words between the anaphoric pronoun and the end of the last sentence. In half of the sentences, N1 was the referent of the anaphoric pronoun (as in Example 1) and in the other half N2 was the referent (as in Example 2). The second set of 12 sentences—all lure— were also adapted from Gernsbacher's (1989) sentences. Names used in the probe task were not used in the stimuli sentences. Lure sentences used the same structure as the 24 experimental sentences (6 with N1 as anaphora referent and 6 with N2 as anaphora referent).

d) Experiment 6

The experimental stimuli were 36 English sentence pairs—24 were experimental sentences (See Appendix D), with the remaining 12 lure sentences. 48 experimental sentence pairs were adopted from Gernsbacher's (1989) Experiment 4. The first sentence in each pair contained two male or female characters, termed N1 (first person mentioned in the sentence) and N2 (second person mentioned in the sentence). One of the characters was referred to in the second sentence by an anaphoric pronoun (either *he* or *she*). The second sentence contained a subordinate clause followed by a main clause which contained the anaphoric pronoun. The subordinate clause was an adverbial participle clause that provided facilitating semantic information about the referent of the anaphoric pronoun.

Example 1: Edward (N1) congratulated Donald (N2) on the recent successful deal. After offering the congratulations, he bought a round of drinks.

Example 2: James (N1) tutored Brian (N2) in clinical and behavioural psychology. Never having been very good in psychology, he really enjoyed the tutoring session.

Stimulus sentences were constructed so that there were five words between N2 and the end of the first sentence, and five words between the anaphoric pronoun and the end of the second sentence. In half of the sentences, N1 was the referent of the anaphoric pronoun (as in Example 1) and in the other half N2 was the referent (as in Example 2). The second set of 12 sentence pairs—all lure— were also adapted from Gernsbacher (1989). Names used in the probe task were not used in the stimuli sentences. Lure sentence pairs used the same structure as the 24 experimental sentence pairs (6 with N1 as anaphora referent and 6 with N2 as anaphora referent).

5.3.2.3 Experimental Procedure

Each comprehension experiment consisted of a comprehension task and a probe recognition task. Experimental procedures for the comprehension task are detailed in Chapter 2, Section 2.3.6. Participants were presented with a series of sentences (Experiment 3), sentence pairs (Experiment 6) or paragraphs (Experiments 4 and 5) followed by a comprehension question. While reading the sentences/paragraphs, participants were presented with a probe. General information about the probe recognition task is provided in Chapter 2, Section 2.3.7. Probes were displayed at two different testing points. Experiment 3 had the same probe testing points as in Experiment 1 (see Chapter 3, section 3.3.2.3 for details). Experiments 4 and 5 had the same probe testing points as in Experiment 2 (see Chapter 4, section 4.3.2.3 for details). Details of each testing point for Experiment 6 are as follows:

1. Before-Anaphora testing point: At the Before-Anaphora testing point, probes were presented 150 milliseconds after the offset of the last word of the first sentence (e.g., Jennifer taught Michelle how to draw a horse. *PROBE* Being a good teacher, she made the job seem easy.); or
2. After-Anaphora testing point: At the After-Anaphora testing point, probes were presented 150 milliseconds after the offset of the last word of the second sentence (e.g., Jennifer taught Michelle how to draw a horse. Being a good teacher, she made the job seem easy.*PROBE*).

5.3.3 Data and Statistical Analysis

Two dependant variables were employed: (a) accuracy and (b) response time to the probe recognition task. For (a), two accuracy scores were calculated for each participant—both accuracy in response to the comprehension question and accuracy in completion of the probe recognition task. The data from 5 participants were removed as they exhibited outlier

comprehension accuracy scores (more than 2 SD below the mean). The participants' working memory spans ranged from 42 to 75 (*Mean = 58.09, Median = 56.00, SD = 10*).

In total, data from 35 out of 40 participants were included in the full analysis. Statistical information about participants' mean scores in comprehension and probe recognition tasks in each experiment is provided in Table 5.2. For analysis of the RTs to the probe recognition task, only experimental trials that were responded to correctly (in both probe recognition and comprehension accuracy tasks) and with raw RTs greater than 300 milliseconds were included. The remaining data was then further trimmed to exclude raw RTs larger than 2SD from the mean of all the correct experimental trials in each experiment separately (1588.95, 1630.87, 1585.84, and 1507.52 milliseconds in Experiments 3, 4, 5, and 6 respectively). These cut-offs removed 3.6%, 3.8%, 4.7%, and 4.6% of the data from Experiments 3, 4, 5, and 6 respectively. Logarithmic transformation was used to normalize the RT distribution. Mixed effects modelling (MEM) was used to analyse response times.

Table 5.2: Participants' mean accuracies in the comprehension and probe recognition tasks in Experiments 3-6.

(%) Accuracy	**Experiment**	**Max**	**Min**	**Median**	**Mean**	**SD**
Comprehension	3	98.00	70.00	86.00	86.06	0.07
	4	100	55.56	80.56	80.08	0.11
	5	100	63.89	88.89	87.54	0.09
	6	100	55.56	80.56	80.56	0.11
Probe Recognition	3	100	84.00	94.00	93.94	0.05
	4	100	77.78	94.44	91.67	0.06
	5	100	80.56	94.44	93.41	0.05

6	100	88.89	97.22	96.27	0.03

5.4 RESULTS

5.4.1 Comprehension Accuracy

A series of binomial mixed effects models were used to analyse comprehenders' accuracy of responses to the comprehension questions. The effects examined in the analysis included fixed effects of the following factors: participant's gender, participant's education, gender of the sentence characters, trial number, working memory capacity, working memory demands (low or high), timeframe for disambiguation (before anaphora (BA) or after anaphora (AA)), the position of the referent in the sentence (first or second), the testing point for the probe recognition task (before reading the anaphora or end of the sentence), log-transformed RT of the probe recognition task in the preceding trial, and the accuracy of response to both the preceding comprehension question and probe recognition task. All relevant interactions between these fixed effects were also tested. Random effects for participant and stimuli were also included. Including the random slopes in this model did not affect the significant results.

A single model was fitted to the data collected in all experiments to analyse the effects of working memory load and the timeframe for presenting contextual information on comprehension accuracy. Model fitting was performed in a backward-stepwise fashion, followed by forward fitting of maximal random effects structure. Models were evaluated by model fitness comparisons using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The final model for comprehension accuracy is presented in Table 5.3. Random effects included in the final model for comprehension accuracy are presented in Table 5.4.

Table 5.3: *Coefficients and p values of the binomial mixed effects model for comprehension accuracy in the four experiments of chapter 5, with participant and stimulus as random effects.*

Effects	Coefficient	Standard Error	z value	P value
(Intercept)	0.115	0.589	0.195	0.845
Disambiguation : BA	-0.467	0.127	-3.676	< 0.001***
WM Capacity	0.029	0.010	2.883	0.004**
Accuracy in PCT	0.369	0.096	3.851	< 0.001***

*Note: WM = working memory, Disambiguation = timeframe for disambiguation information, PCT= preceding comprehension trial, BA: before anaphora, * = p < 0.05, ** = p < 0.01, *** = p < 0.001*

Table 5.4: *Random effects included in the final binomial mixed effects model for comprehension accuracy in the four experiments of Chapter 5.*

Random effects		Variance	Standard Deviation
Stimulus	(Intercept)	0.391	0.625
Participant	(Intercept)	0.205	0.453

The model revealed a significant main effect of the timeframe for disambiguation *(p < 0.001)*, indicating that comprehension accuracy was significantly lower in the experiments in which contextual information required for anaphora resolution was provided before anaphora (Experiments 4, and 6) compared to after anaphora (Experiments 3, and 5). The

main effect of tasks' working memory storage demand was, however, not significant and was removed from the model.

The main effect of working memory capacity was also significant *($p < 0.01$)*, demonstrating that as working memory capacity increased, so did the comprehension accuracy. Moreover, performance on the preceding comprehension task ($p < 0.001$) also significantly affected performance on the comprehension task. Participants were more accurate in answering comprehension questions after correctly answering the preceding comprehension trial.

5.4.2 Response Times

For each experiment, a series of linear mixed effects models were used to analyse participants' response times (log-transformed RTs) to the probe recognition task. The analysis examined the fixed effects of participants' gender, participants' years of education, working memory capacity (Reading Span score), probe type (referent (REF) or nonreferent (NREF)), probe testing point (before anaphora or at the end of the sentence), referent position in the sentences (first or second), probe position in the sentences, trial number, log-transformed RT to the probe recognition task in the preceding trial, and the accuracy of responses to both the preceding comprehension question and probe recognition task. All relevant interactions between fixed effects were tested. Random effects for participant and stimulus were also included.

Model fitting was performed in a backward-stepwise iterative fashion, followed by forward fitting of maximal random effects structure. Model fitting was independently supported by model fitness comparisons using AIC and BIC.

5.4.2.1 Experiment 3

The final model for log-transformed RTs in Experiment 3 is presented in Table 5.5. This model included random intercepts for stimuli and participants, as well as by-participant random slopes for trial number, and RT to the preceding probe recognition task. The random intercepts and by-participant random slopes included in the final model are presented in Table 5.6. The model revealed two significant main effects of trial number ($p < 0.001$) and response time in the preceding probe recognition trial ($p < 0.01$). Participants got faster in responding to the probes with the increasing number of trials and their response times were positively affected by their response times in the preceding trial.

Table 5.5: *Coefficients and p values of the final linear mixed effects model for RTs in Experiment 3, including fullest random effects structure.*

Effects	**Coefficient**	**Standard Error**	**t value**	**P value**
(Intercept)	6.806	0.038	181.180	
Trial Number	-0.022	0.005	-4.430	< 0.001***
RT in PRT	<0.001	<0.001	2.800	0.005**

*Note: PRT = preceding recognition trial, * = p < 0.05, ** = p < 0.01, *** = p < 0.001*

Table 5.6: *Random effects and slopes included in the final linear mixed effects model for RTs in Experiment 3.*

Random effects		**Variance**	**Standard Deviation**
Stimulus	(Intercept)	<0.001	<0.001
Participant	(Intercept)	0.027	0.164
	Trial Number	<0.001	<0.001

RT in PRT	<0.001	<0.001
Residual	0.028	0.167

Note: PRT = preceding recognition trial

5.4.2.2 Experiment 4

The final model for log-transformed RTs in Experiment 4 is presented in Table 5.7. This model included random intercepts for stimuli and participants, as well as by-participant random slopes for trial number, probe recognition testing point, probe type, and the interaction of probe recognition testing point and probe type. The random intercepts and by-participant random slopes included in the final model are presented in Table 5.8.

Table 5.7: *Coefficients and p values of the final linear mixed effects model for RTs in Experiment 4, including fullest random effects structure.*

Effects	**Coefficient**	**Standard Error**	**t value**	**P value**
(Intercept)	6.893	0.021	327.6	
Trial Number	-0.049	0.012	-4. 0	< 0.001***
PTP: END	0.093	0.021	4.4	< 0.001***
Probe Type: REF	0.005	0.020	-0.2	0.841
PTP: END*Probe Type: REF	-0.115	0.031	-3.7	< 0.001***

*Note: PTP = probe testing point, REF = referent, END = end of the sentence, * = p < 0.05, ** = p < 0.01, *** = p < 0.001*

Table 5.8: *Random effects and slopes included in the final linear mixed effects model for RTs in Experiment 4.*

Random effects		**Variance**	**Standard Deviation**
Stimulus	(Intercept)	< 0.001	< 0.001
Participant	(Intercept)	0.008	0.089
	Trial Number	0.003	0.051
	PTP: END	0.002	0.041
	Probe Type: REF	< 0.001	0.021
	PTP: END*Probe Type: REF	0.005	0.075
Residual		0.029	0.169

Note: PTP = probe testing point, REF = referent, END = end of the sentence

As expected, trial number *($p < 0.001$)* had significant effects on the response time. Participants got faster in responding to the probes with the increasing number of trials. Another significant effect was the main effect of probe recognition testing point *($p < 0.001$)* showing that responses were significantly slower at the end of the sentence compared to the before anaphora testing point. In addition, this model revealed a significant two-way interaction between probe type and probe recognition testing point *($p < 0.001$)*. Responses to the nonreferent probes were slower compared to the referent probes at the end of the sentence. This interaction is depicted in *Figure 5.1*.

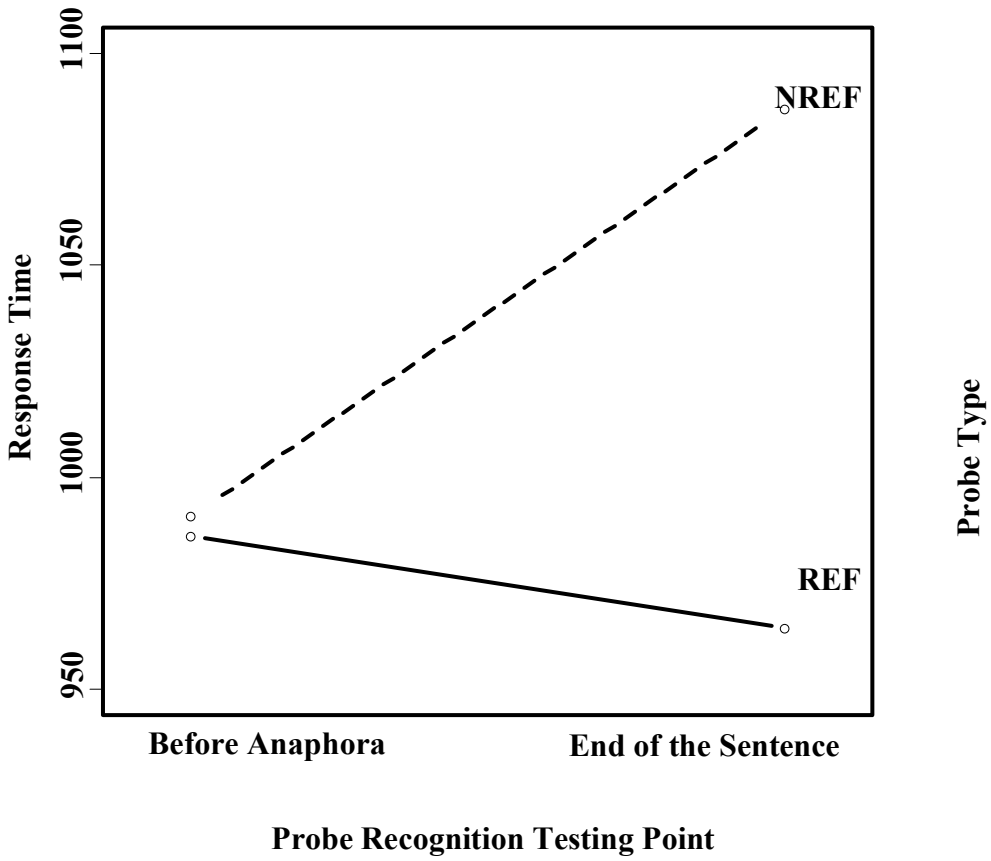

Figure 5.1: Two-way interaction of probe type and probe recognition testing point

5.4.2.3 Experiment 5

The final model for log-transformed RTs in Experiment 5 is presented in Table 5.9. This model included random intercepts for stimuli and participants. The random intercepts included in the final model are presented in Table 5.10. The model only revealed one significant main effect of response time in the preceding probe recognition trial ($p < 0.001$). Participants' response times were positively affected by their response times in the preceding trial.

Table 5.9: *Coefficients and p values of the final linear mixed effects model for RTs in Experiment 5, including fullest random effects structure.*

Effects	Coefficient	Standard Error	t value	P value
(Intercept)	6.708	0.038	175.290	
RT in PRT	<0.001	<0.001	3.760	<0.001**

*Note: PRT = preceding recognition trial, * = p < 0.05, ** = p < 0.01, *** = p < 0.001*

Table 5.10: *Random effects and slopes included in the final linear mixed effects model for RTs in Experiment 5.*

Random effects		Variance	Standard Deviation
Stimulus	(Intercept)	0.004	0.066
Participant	(Intercept)	0.016	0.0125
Residual		0.034	0.185

5.4.2.4 Experiment 6

The final model for log-transformed RTs in Experiment 6 is presented in Table 5.11. As in the model for the analysis in Experiment 6, this model included random intercepts for stimuli and participants, as well as by-participant random slopes for trial number, probe recognition testing point, probe type, and the interaction of probe recognition testing point and probe type. The random intercepts and by-participant random slopes included in the final model are presented in Table 5.12.

Table 5.11: *Coefficients and p values of the final linear mixed effects model for RTs in Experiment 6, including fullest random effects structure.*

Effects	**Coefficient**	**Standard Error**	**t value**	**P value**
(Intercept)	6.749	0.026	262.24	
Trial Number	-0.038	0.013	-2.96	0.003**
PTP: END	0.049	0.024	2.00	0.045*
Probe Type: REF	0.002	0.024	0.10	0.920
PTP: END*Probe Type: REF	-0.075	0.037	-1.99	< 0.046*

*Note: PTP = probe testing point, REF = referent, END = end of the sentence, * = p < 0.05, ** = p < 0.01, *** = p < 0.001*

Table 5.12: *Random effects and slopes included in the final linear mixed effects model for RTs in Experiment 6.*

Random effects		**Variance**	**Standard Deviation**
Stimulus	(Intercept)	< 0.001	0.022
Participant	(Intercept)	0.013	0.115
	Trial Number	0.003	0.054
	PTP: END	< 0.001	0.030
	Probe Type: REF	< 0.001	0.028
	PTP: END*Probe Type: REF	0.008	0.089
Residual		0.032	0.178

Note: PTP = probe testing point, REF = referent, END = end of the sentence

The results of this experiment replicated the results from Experiment 4. The model revealed a significant main effect of probe recognition testing point *(p < 0.05)* showing that responses were significantly slower at the end of the sentence compared to before anaphora

testing point. The two-way interaction of probe type and probe recognition testing point also reached significance *(p < 0.05)*. Responses to the nonreferent probes were slower compared to the referent probes at the end of the sentence. This interaction is depicted in *Figure 5.2*. Trial number *(p < 0.01)* also significantly affected response times demonstrating that participants got faster in responding to the probes with the increasing number of trials.

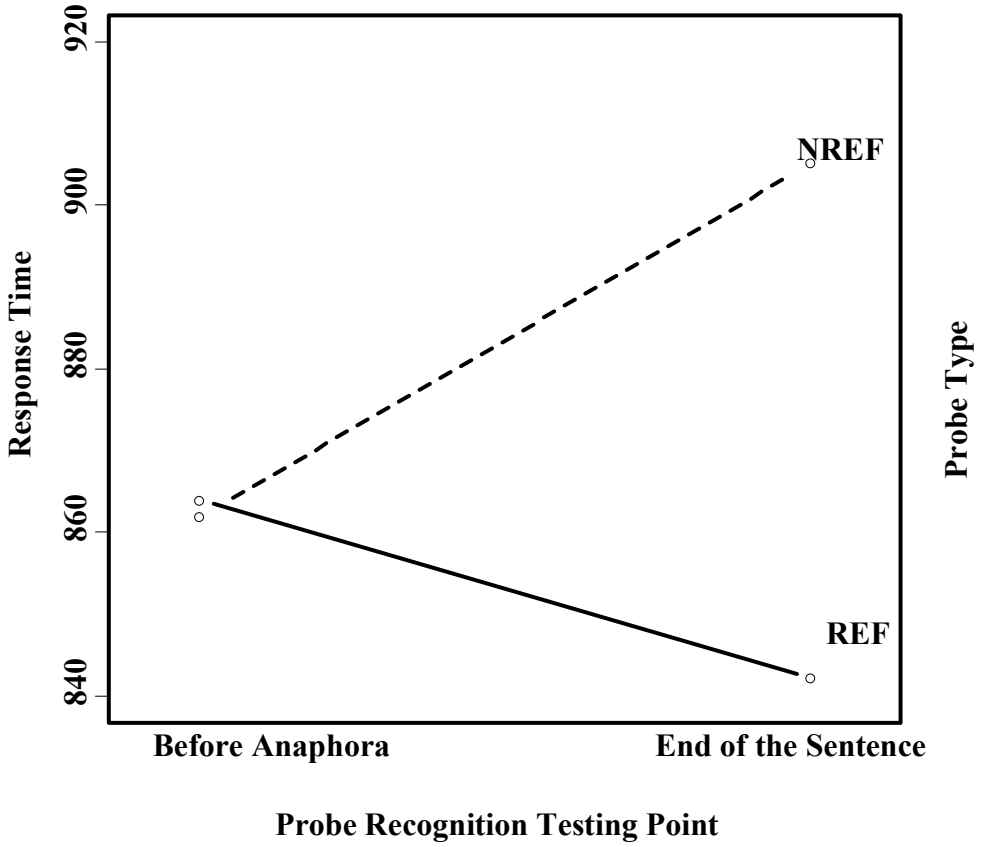

Figure 5.2: Two-way interaction of probe type and probe recognition testing point

5.5 DISCUSSION

The current study aimed to determine whether earlier anaphora resolution in Experiment 2 (Chapter 4) was due to increased working memory storage load, high processing load or earlier disambiguation. Results indicated that increased processing demands of the more syntactically complex sentences contributed to earlier anaphora resolution. Overall, this was interpreted to provide further support for the concept that anaphora resolution was strategic rather than automatic and depended on task demands.

Experiment 3 had low processing load, low storage load and late disambiguation. All these factors were manipulated in Experiment 4, which had high processing load, high storage load and prior disambiguation. Experiments 5 and 6 aimed to separately examine the effects of the manipulation of these factors on the time-course for anaphora resolution. Experiment 5 was used to investigate the effects of increased storage demands, and Experiment 6 was to demonstrate the effects of increased processing demands as well as prior disambiguation. Results indicated that participants suppressed the accessibility of nonrefernts by the time they finished reading the sentences in Experiments 4 and 6. However, suppression was not observed in Experiments 3 and 5. This suggests that in Experiments 4 and 6, anaphora resolution was completed by suppressing the nonreferent name by the end of the sentence while it was delayed in the other two experiments.

As in Chapters 3 and 4, response times were used to study whether increased storage load, increased processing load or prior disambiguation would result in earlier suppression of the nonreferent. The response time results from Experiment 3 replicated the finding from Chapter 3 of this study (Experiment 1, low working memory storage load). It showed delayed anaphora resolution under low working memory storage load, low processing load, and in the absence of prior disambiguation. This finding provided further support for the hypothesis that

under low working memory demands, readers might delay the process of anaphora resolution because their working memory capacity is sufficient for maintaining and processing of all the information.

Results from Experiment 4 also replicated the finding from chapter 4 of this study (Experiment 2). It demonstrated earlier anaphora resolution under high working memory storage load, high processing load and in the presence of prior disambiguation. Since three factors were manipulated in this experiment compared to the first experiment, it is unclear which of these factors contributed to earlier resolution of anaphora. Therefore, the effect of increased storage load was separately examined in Experiment 5 which had low processing demands and did not provide prior disambiguation.

As in Experiment 3, Results from Experiment 5 suggested delayed resolution of anaphora. Moreover, comprehension accuracy data showed that working memory storage load did not affect comprehension accuracy. If the earlier suppression effect observed in Experiment 4 were due to higher working memory storage demands, we would expect the accessibility of nonreferents to be suppressed by the time readers finished reading the sentences in Experiment 5. However, the lack of suppression in Experiment 5 suggested that the suppression effect observed in Experiment 2 (see Chapter 4) and Experiment 4 in the current chapter was not due to higher working memory storage demands. Rather, it might have been triggered by either higher syntactic complexity or facilitating effect of prior disambiguation. Therefore, the effects of these factors on anaphora resolution was examined in Experiment 6.

Results from Experiment 6 demonstrated that participants resolved anaphora by the time they reached the end of the sentence under low working memory storage load when disambiguating information was provided in an adverbial phrase preceding the anaphora.

Sentences containing anaphoric pronouns in this experiment were syntactically complex and more demanding than the simple and compound sentences used in Experiments 3 and 5. The finding from this experiment thus suggested that earlier anaphora resolution was due to either prior disambiguation or high processing demands of the task.

If earlier anaphora resolution was due to prior disambiguation, comprehension accuracy was expected to be higher in this experiment, as prior disambiguation was to facilitate comprehension. However, comprehension accuracy results indicated that participants had significantly lower scores on this experiment compared to Experiments 3 and 5. This finding suggested that prior disambiguation, through adding a subordinate clause to the anaphoric sentences in this experiment, was not facilitating but rather made the anaphoric sentences more difficult to process. Therefore, earlier suppression cannot be attributed to a facilitation effect caused by prior disambiguation. Higher processing demands of syntactically complex sentences might have contributed to the earlier suppression of nonreferents during anaphora resolution. This finding can be well explained by the shared-capacity theory of working memory proposed by Daneman and Carpenter (1980).

In this theory, working memory capacity was considered a cognitive resource shared among storage and processing. This theory suggested that there was a trade-off between storage and processing and the processing demands received priority. Therefore, under high processing load, storage would be deficient. This trade-off between storage and processing was observed in the current study. Participants in this study suppressed the information from working memory storage to devote larger capacity to processing under higher processing load. Suppression of the irrelevant discourse information from working memory storage thus might have contributed to improving the processing ability. Our findings, therefore, provide support for the shared-resources and the trade-off theories of working memory capacity.

5.6 CONCLUSION

Study results showed that comprehension accuracy was negatively affected by syntactic complexity of sentences while increased storage load did not influence comprehension accuracy. Moreover, the present study provided further support for the strategic rather than automatic processing of anaphora resolution. Results demonstrated that working memory storage load and prior disambiguation did not affect the process of anaphora resolution. In contrast, anaphora resolution was faster under high processing load. It was suggested that under lower working memory processing demands, younger adults might postpone the resolution of anaphora. Higher processing demands, however, might require earlier resolution of anaphora so that irrelevant information could be suppressed earlier from working memory to provide larger capacity for processing. Therefore, suppression of irrelevant discourse information was suggested to contribute to language comprehension abilities through improving processing abilities. The current study findings provided support for a shared-resources theory of working memory showing a trade-off between processing and storage.

6. CHAPTER SIX

Summary, Limitations, and Future Directions

6.1 SUMMARY

The research described in this thesis provides an insight into the cognitive mechanisms underlying language comprehension and the contributions of the age-related decline in more general cognitive functions to language comprehension difficulties. Background information provided in Chapter 1 acknowledged that some cognitive abilities including high-level language comprehension declines with ageing and that the nature and the underlying mechanisms of comprehension decline in ageing have remained controversial.

Considering the important roles of working memory functions in high-level language processing, and the negative effects of age on working memory functions, Chapter 1 raised the possibility that age-related decline in working memory abilities might contribute to changes in language comprehension. A review of the literature revealed some evidence for older adults' difficulties in comprehending potentially ambiguous discourse meanings such as finding the referents of discourse anaphora.

Anaphora resolution was recognized as a discourse processing skill which has been commonly reported to be affected by ageing. However, relatively few studies have investigated the effects of age and working memory abilities on anaphora resolution. Furthermore, most of these studies have used offline paradigms which could not address the contributions of working memory functions to the online process of reference resolution. Chapter 1 highlighted the need for an online study of reference resolution in ageing, in a framework that allows for the investigation of the underlying cognitive processing involved. Being able to account for the individual differences in cognitive abilities, SBF was introduced as a good framework in which to study the cognitive processing involved in anaphora resolution.

Chapter 2 detailed a general description of the methodology employed across three experimental chapters (Chapters 3, 4, and 5) of the current research. Chapters 3 and 4 included two comprehension experiments and an assessment of working memory capacity, and Chapter 5 consisted of four comprehension experiments and an assessment of working memory capacity. Gernsbacher's (1989) paradigm was adapted to investigate a) the accuracy in finding the referents of potentially ambiguous anaphoric pronouns, and b) the accessibility of the referents and nonreferent names before and after anaphora resolution. The accessibility of the names was assessed through measuring response times in a probe recognition task presented to the participants before or after reading the anaphoric pronoun in a series of sentences. Working memory storage load, processing load, and the time-course for presenting disambiguating contextual information were manipulated across experiments. Binomial and linear mixed effects modelling were used to analyse comprehension accuracy and response times respectively.

Chapter 3 reports the findings from the first experiment of the research (Experiment 1, low working memory storage load). Results revealed that comprehension accuracy was positively affected by working memory span and negatively influenced by age. Findings from recognition task showed that working memory capacity, but not age, affected the process of anaphora resolution. First mentioned names were more accessible for participants with higher working memory. This finding was suggested to be the result of laying the foundation for mental representation. In contrast, recently mentioned names were more accessible for participants with lower working memory capacity suggesting that they relied on recency instead of building a mental representation through laying a foundation. Moreover, this experiment provided evidence that suppression of nonreferents and thus anaphora resolution might be delayed under lower working memory load. From the previous studies on anaphora resolution, it was expected that the process of anaphora resolution be completed by the time

participants completed reading the sentences. As a result, the accessibility of the nonreferent names was expected to be suppressed by the time readers reached the end of the sentence. Results from this experiment, however, did not demonstrate suppression of the nonreferents either in the younger or in the older adults. Although older adults were less accurate than younger adults in resolving anaphora, no effect of age was found on the process of anaphora resolution. To further investigate the effects of age on the process of anaphora resolution, a second experiment was used which had higher working memory demands.

Chapter 4 presented the findings from the second experiment of the research (Experiment 2). Here, working memory storage demand was increased to examine whether higher working memory load might affect anaphora resolution and result in earlier suppression of nonreferents. To increase the task's working memory demands, referential distance was increased through adding two intervening sentences between the clause introducing the characters and the clause containing the anaphoric pronoun. As in the Chapter 3, age and working memory capacity affected comprehension accuracy. Moreover, the results provided further evidence for the previously found effect of working memory capacity on the process of anaphora resolution. Results also revealed that age affected suppression processing during the online resolution of anaphora. Younger adults suppressed the nonreferents by the time they reached the end of the sentence, while older adults showed no suppression. It was concluded that comprehension decline in older adults might be mediated by a suppression deficit. This finding partially supported the inhibitory decline theory of cognitive ageing.

The finding from Chapters 3 and 4 showing that suppression processing during anaphora resolution was delayed in the Experiment 1 (low working memory storage load , See Chapter 3), but not in the Experiment 2 (high working memory storage load, See Chapter 4), established the rationale for the next part of the research outlined in Chapter 5. Chapter 5

further investigated the cognitive mechanisms involved in anaphora resolution to find the factors contributing to the earlier anaphora resolution in Experiment 2 (See Chapter 4), compared to in Experiment 1 (See Chapter 3),. In Chapter 5, participants completed four comprehension experiments and an assessment of working memory capacity. Referential distance and syntactic structure of anaphoric sentences were manipulated. Referential distance was increased to examine the influence of working memory storage load on anaphora resolution. Syntactic structure of sentences was manipulated to examine the effects of syntactic complexity and prior disambiguation on anaphora resolution.

Working memory storage load did not affect accuracy of anaphora resolution. Nor did it affect the process of anaphora resolution. In contrast, changing the syntactic structure of anaphoric sentences resulted in earlier suppression of the nonreferents—earlier anaphora resolution. Suppression was observed only in anaphoric sentences that provided disambiguating information in a dependant clause prior to the anaphoric pronoun. Since comprehension accuracy was lower in these sentences, it was concluded that the higher processing load of these syntactically complex sentences, rather than the facilitating effect of prior disambiguation, resulted in earlier suppression of the nonreferents. The results were explained using the shared resources theory of working memory capacity. It was suggested that under higher processing demands, a trade-off between working memory storage and processing might have resulted in an earlier suppression of information from working memory storage so that larger capacity was available for processing.

Collectively, studies reported in chapters 3, 4, and 5 revealed that anaphora resolution might be strategic rather than automatic. When processing demands were low, limited working memory capacity was enough for simultaneous processing and storage of the information. Therefore, the irrelevant information was not required to be suppressed and

anaphora resolution could be delayed. When processing demands of the task were higher, the limited capacity of working memory was exceeded. Therefore, for successful comprehension, the irrelevant information needed to be suppressed earlier from working memory. In such a case, anaphora was resolved earlier. Our findings demonstrated that while younger adults suppressed the irrelevant information under higher processing load, older adults were unable to suppress that information. Since older adults were less accurate than younger adults in answering comprehension questions regardless of the task's working memory demands, it was suggested that their comprehension decline was due to inefficient suppression skills.

Overall, the present thesis aimed to increase understanding of the process of anaphora resolution, and the contributions of age-related differences in cognitive abilities to comprehension abilities. A better knowledge of older adults' comprehension abilities can help in keeping them involved in social interactions. This study suggests that the use of potentially ambiguous linguistic structures should be limited in conversations with older adults and in their media. Moreover, minimizing the use of these structures in written materials aimed at older adults, including healthcare brochures, safety instructions, and medicine information sheets, can contribute to their health and safety. Furthermore, the current findings can be applied to the research on pathologically aged and neurologically impaired older adults. High-level language comprehension has been mostly reported to be impaired in such populations. It is possible that their high-level comprehension impairments are the results of normal ageing rather than due to a neurological lesion. In order to understand language impairments caused by a neurological disease, it is crucial to have knowledge of language abilities in normal ageing.

6.2 LIMITATIONS AND FURTHER DIRECTIONS

The studies described in the current thesis are limited by a number of factors. These limitations have implications for the future direction of research into comprehension abilities in ageing. Limitations in this study are discussed in separate sections below.

6.2.1 Participants

The primary limitation in this study was the number of participants. Recruiting older adults over 65 who were both willing and able to participate in the research was not an easy task. Therefore, the number of participants was limited due to the limited time available for recruitment and limited accessibility to older adults. Moreover, older adults were generally less accurate than younger adults. Therefore, a number of the older participants had to be excluded from the final analysis to make sure that response times were not based on chance and random guesses. Although dataset imputation revealed that a larger dataset would not change the significant results, including more participants in the study can increase the statistical power and make it possible to include four-way interactions in the statistical model.

Furthermore, the exclusion criteria used in the previous research with a similar paradigm could not be used in this research. The previous research included only the younger participants and participants were replaced if their accuracies were less than 85% in answering the comprehension questions or less than 90% in answering the probe recognition task. These exclusion criteria would exclude the majority of older adults from the study. Therefore, a more conservative exclusion criteria was used in the current study to better suit the abilities of the older participants.

6.2.2 Comprehension Experiment

Although the paradigm used for the comprehension experiment in this study was suggested to best suit the current study aims, it had some limitations that can be addressed in

future comprehension studies. First, the presentation of the recognition task while participants were performing the reading comprehension task had the potential to disrupt the process of comprehension and cause distraction. Results also demonstrated that comprehension accuracy was affected by testing point for the recognition task, particularly in less demanding tasks. Moreover, this distraction might have affected older adults more than younger adults as they have been reported in some studies to have difficulty performing dual tasks requiring divided attention (Allen, Ledgeway, Kelly, Hutchinson, & Blundell, 2011; Park, Denise, Anderson, William, & Vincent, 1989).

Second, the probe recognition task might have resulted in the use of a task-specific strategy by participants to improve their memory for discourse details, particularly name of the characters. Therefore, participants, particularly those with smaller working memory capacity, might have not been fully engaged in discourse comprehension. Third, although users of language, as skilled listeners, are experienced in sequential processing of language, it might be argued that a time-controlled word-by-word presentation of text in the current study disrupted the natural process of reading comprehension. Use of different text presentation methods could help investigating whether reading comprehension might be affected by text presentation methods. Therefore, since comprehension strategies are argued to be influenced by the specific requirements of the tasks used for assessment, it is commonly suggested that different paradigms be used to examine comprehension process before drawing definite conclusions. Replicating the findings using different paradigms can reduce the possibility that the results were affected by task-specific demands and presentation method.

6.2.3 Sentence Stimuli

Although sentences used in the experiment were controlled for the number of words after the pronoun anaphora, and between the second name and pronoun anaphora, some

structural differences existed in these sentences. The first sentences, which introduced the discourse characters, were simple in some trials and complex in the others. Therefore, since sentence complexity affects processing demands, these sentences might have had different processing demands. This might have caused some differences in the processing of sentences across different trials and thus affected the results.

Moreover, the two character names were not always included in the same sentence. In a number of complex sentences, the second mentioned character was introduced in an imbedded sentence while the first character was mentioned in the main clause. Although Gernsbacher had demonstrated that the effect of clause recency was short-lived and did not last longer than the end of the clause (Gernsbacher et al., 1989), it might be possible that accessibility of the names was affected by their placement in different clauses. Controlling for the structure of sentences could eliminate the possible effects of sentence complexity as well as clause recency.

6.2.4 Measure of Inhibitory Abilities

In this study, only one measure of working memory abilities was used. Since an aim of this study was to investigate the influence of individual differences in working memory capacity on anaphora resolution, an assessment of working memory capacity was included. A reading span task was used for assessment as it has been commonly suggested to better correlate with comprehension scores. Scores on the reading span task could predict some individual differences in anaphora resolution. However, they failed to predict age-related changes in anaphora resolution in the current study. Age-related changes in this study existed in suppressing the irrelevant discourse information and were hypothesized to be due to inefficient inhibitory functions. This hypothesis could not be further examined in this study, as the study did not include any measure of inhibitory abilities.

Reading span task is mainly a measure of the ability to store, and maintain information under distracting condition through attentional control. Although performing the reading span task involves inhibition, the inhibitory functions important to performing the reading span task are suggested to be different from those that are crucial to discourse comprehension. Since different inhibitory functions are suggested to exist, using different measures of inhibition, including measures of comprehension-related inhibition, can increase understanding of the contributions of inhibitory decline to age-related changes in discourse comprehension. Furthermore, including different measures of inhibitory functions could help examine the hypothesis that different inhibitory abilities might be influenced differently by normal ageing. Therefore, further investigation of how the accuracy and the process of anaphora resolution might be affected by differences in inhibitory abilities can shed lights on the roles of different inhibitory functions in language comprehension and the effect of age on these functions.

6.3 SUMMARY AND CONCLUSION

Age-related changes in high-level language comprehension and their underlying cognitive mechanisms were identified as important areas for further research. The current study has provided evidence that older adults face difficulties in resolving anaphora, particularly when it is possible to infer more than one potential referent. Moreover, the findings suggested that this language processing difficulty might be caused by an age-related decline in suppression abilities. Furthermore, this thesis has highlighted the contribution of working memory capacity to differences in building the mental representation during discourse comprehension. It is suggested that both differences in working memory capacity and age-related decline in suppression abilities can affect discourse comprehension. Current understanding of the underlying mechanisms involved in language processing and age-related

changes in language comprehension is anticipated to improve with further research including greater numbers of participants, use of different paradigms for comprehension assessment, and measures of inhibitory abilities.

APPENDIX A

Sentence Stimuli for Comprehension task: Experiment 1

Experimental Sentences

1. Katherine stood up until Christina had brought in another chair then she sat down on the chair.
2. Roy wanted to tell Joe the exciting and unexpected news but he couldn't find a nearby phone.
3. Alan sent Cory to do the grocery shopping and he returned with several heavy sacks.
4. Frank loaned Jerry a blue ball point pen but he wanted it back before long.
5. Christian handed Alexander some tickets to a concert but he took the tickets back immediately.
6. Stacey walked Pamela over to the dentist's office but she waited outside in the lobby.
7. Anne borrowed a book from Leah all about World War II but she never gave the book back.
8. Albert saw that Darren was in very serious trouble and he ran quickly for some help.
9. Nathaniel saw Frederick standing on the river bank and he waved hello from the canoe.
10. Michael asked William to play a round of golf but he had already made other plans.
11. Angela gave Nicole some directions to the zoo and she had no trouble following them.
12. Matthew tried to beat Jeffrey in a game of chess but he managed to win every time.
13. Brandy found out that Alicia was feeling a little sick but she made a very speedy recovery.
14. Beth received from Jodi one of those chain letters but she did not continue the chain.
15. Adrian poured a drink for Ronnie that was really quite strong then he poured a drink for himself.
16. David watched Jason act in a Broadway play and he applauded at the final curtain.
17. Danielle predicted that Kathleen would lose the track race but she came in first very easily.
18. Sara asked Erin to pick out a card and she drew the ace of diamonds.
19. Dustin spilled a drink on Gerald at the New Year's party and he went home to change clothes.
20. Tara sent Lori a cheque for twenty dollars and she cashed the $20 cheque immediately.
21. Felicia waited for Colleen in the fancy restaurant lounge and she arrived a half hour late.
22. Kristina accused Veronica of committing a big robbery and she was convicted of the crime.
23. Lee saved Jay from drowning in the creek and he quickly became a big hero.
24. Carl expected Luis to arrive on the train but he was not on the train.
25. Natalie loaned Valerie an interesting book on astronomy after she finished reading the book herself.

26. Stephanie went to visit Elizabeth during the hospital's visiting hours and she brought a bouquet of flowers.
27. Cheryl wanted a snapshot of Leslie in front of the museum but she wouldn't pose for the camera.
28. Craig loaned Peter some tools for the garden and he returned them a week later.
29. Monica tutored Sandra in history, math, and English and she charged ten dollars an hour.
30. Shane blamed Larry for causing the car accident but he was really not at fault.
31. Curtis pitched Victor a very fast curve ball and he hit it into the outfield.
32. Andrea gave Amanda some truly heartfelt advice yesterday but she didn't take the advice seriously.
33. Lawrence broke a leg while skiing with Nicholas at a very expensive resort and he had to leave on crutches.
34. Seth made sure that Toby was already very sound asleep and he tiptoed out of the house.
35. Jonathan saw that Benjamin was fixing a flat tyre and he stopped to offer some help.
36. Kyle scratched Erik with a pocket knife accidentally and he started bleeding from the wound.
37. Joel threw a pie at Tony that was big and gooey but he ducked before it could hit.
38. Carrie mailed Rachel a package of secret information and she received it within a week.
39. Anthony tied Charles to a big wooden chair but he was able to get loose.
40. Brian was amusing James by doing some fancy acrobatics but he slipped and broke an arm.
41. Bryan tried to amuse Keith with a somewhat off-colour joke but he didn't even laugh at it.
42. Gina made Anna a rich chocolate pound cake and she used an old fashioned recipe.
43. Carolyn lost to Yolanda in the national tennis match but she accepted the major defeat gracefully.
44. Karen gave Tracy a very long boring lecture and she listened to it very patiently.
45. Brenda bought a car from Sharon that was eight years old and she was pleased with its performance.
46. Philip locked Johnny out of the house accidentally and he broke in through a window.
47. Richard passed the ball to Timothy in a game of soccer but he lost the ball very quickly.
48. Dana described to Jill how life was in Sydney but she didn't mention the terrible pollution.
49. Tina cleaned the house for Dawn for several hours one day while she napped peacefully on the sofa.
50. Willie saw Wesley outside stealing a parked car but he did not call the police.
51. Jack handed Brad the telephone in the den after he had gotten tired of talking.
52. Margaret called Victoria on a special phone line and she answered on the third ring.
53. Tammy thought that Laura was hard at work studying but she had gone to a movie.
54. Walter aimed a pistol at Martin that looked like a toy but he did not pull the trigger.
55. Teresa was knitting a scarf for Denise for an early Christmas present but she did not have enough yarn.
56. Katrina was being tickled by Sabrina while they were watching TV but she managed not to laugh aloud.
57. Carlos punched Rodney during a bar room brawl and he got a terrible black eye.

58. Jennifer borrowed a bike from Michelle to go to the university and she had to return it quickly.
59. Patricia inherited from Kimberly a very substantially large fortune and she spent all the money foolishly.
60. Courtney urged Samantha to apply to law school and she got accepted the next year.
61. Mary went to visit Lisa one rainy afternoon in July but she was away on a vacation.
62. Shawn begged Aaron to play a game of handball and he reluctantly agreed to do that.
63. Vanessa interviewed Suzanne about cheating in college courses but she refused to answer some questions.
64. Kevin wrapped a gift for Scott that was a big surprise and he hid it in the closet.

APPENDIX B

Sentence Stimuli for Comprehension task: Experiment 2

Experimental Paragraphs

1. Angela sang a song for Nicole which was old but popular. It was a classic love song. It had been first performed in 1970s. Carefully listening to the words, she wanted to cherish their meaning.
2. Victor got a postcard from Curtis with a picture of beach. The sandy beach in the picture was full of shells. That made it more beautiful. Though jealous about the vacation, he enjoyed staying in touch with friends.
3. Valerie expected Natalie to arrive early for dinner. The dinner was supposed to be served at 8:30pm. Even the dessert was ready. After waiting for over an hour, she was ready for the guests.
4. Carlos borrowed some money from Rodney to buy an expensive watch. It was a beautiful golden watch. It was on sale for a limited time. Grateful for the loan, he felt a sense of comradely.
5. Brian got a letter from James along with twenty five photos. The photos were very beautiful. They pictured the beautiful scenery of Mexico. While vacationing in Mexico, he enjoyed staying in touch with family.
6. David gave Jason a ride to the library. It was a big library full of academic books. There were also lots of non-academic materials there. While parking the car, he was thinking about a book.
7. Rachel drew a picture of Carrie sitting in a beautiful park. The scenery was amazing. The park was full of trees and flowers. After posing for several hours, she was pleased with the drawing.
8. Darren rescued Albert from drowning in the pool. The pool was very deep. It was a private pool without lifeguards. Enjoying being a hero, he talked about it for hours.
9. Wesley pushed Willie into the outdoor swimming pool. It was a really cold winter day. The water in the swimming pool was freezing. After doing such a mean thing, he ran away quickly to hide.
10. Beth aimed a water gun at Jodi while playing in the school. It was a nice long shot water gun. It was made in the USA. Ready to shoot, she thought of a better idea.
11. Frank mowed the front lawn for Jerry after coming back from school. It was a nice summer day. It was very sunny outside. After finishing the mowing, he trimmed all of the hedges.
12. Lee lost a book that belonged to Roy in the park last night. It was an expensive academic book. It was very old and really rare. After apologizing, he offered to buy another one.
13. Bryan taught Keith how to sing a song. It was a very beautiful love song. The song was very famous among the teenagers. Being a good learner, he made the job seem easy.
14. Jonathan locked Benjamin out of the flat yesterday. It was a winter night. It was very dark and cold outside. After realizing the mistake, he put a key under the mat.

15. Colleen found a pen that belonged to Felicia on the old wooden table. It was one of those multicolour pens. It had four different colours. After realizing that it was missing, she looked around for another pen.
16. Sandra mailed a letter to Monica using the international express service. The letter was a job offer. It contained some information about the job and the salary. Sending the letter first class, she hoped it would arrive safely.
17. Ronnie beat Adrian in an entertaining video game. It was a famous fighting game. The game was very violent. Being a horrible winner, he talked about the game endlessly.
18. Walter always read books to Martin in the school break holidays. The books were usually about tourism. They had some information about different countries. Though hating to read out loud, he liked knowing about other countries.
19. Jeffrey broke a glass that belonged to Matthew while helping in the kitchen. It was a nice antique glass. It was a beautiful blue colour. After saying not to worry about it, he looked around for the broom.
20. Shawn handed the telephone to Aaron to call the hotel reception. It was the best hotel in the city. Therefore, it was always fully booked. After picking up the receiver, he sat down on a chair.
21. Seth read Dale a story about a kitten. The cute kitten in the story was very playful. It loved playing with knitting balls. After finishing the story, he decided to buy a kitten.
22. Alicia called the firemen to save Brandy from a burning university building. The fire was started by an oven explosion. It spread very quickly to all buildings. Waiting to be rescued, she was eager for their arrival.
23. Katrina cooked Sabrina a nice meal for breakfast. The meal was extremely large. There was lots of food on the table. While preparing the huge meal, she hoped everyone was hungry enough.
24. Alan stole the basketball from Luis in a very competitive game. There were some valuable prizes for the winning teams. The best prize was a trip to Sydney. After losing the ball, he heard the coach yelling angrily.
25. Dawn greeted Tina with smiles in the airport. The airport was very busy. A number of flights had landed. Surprised by the warm welcome, she began to get teary eyed.
26. Gerald read the local newspaper to Dustin every day during afternoon tea. It was always full of ads. However, it had a special column on world news. Having been blind since birth, he liked knowing about current events.
27. Karen repeated the story for Tracy about the last month's robbery. Robbers had broken into a jewellery shop. An expensive diamond ring had been stolen. Not having heard clearly the first time, she tried even harder to concentrate.
28. Veronica mailed a package to Kristina containing a pair of shoes. The shoes were very expensive. They were made by a famous shoemaker. Eager to receive the package, she hoped it would arrive quickly.
29. Cheryl helped Leslie across the very wide stream. It was a very beautiful stream. However, it was very deep. After being kindly assisted, she looked back across the stream.
30. Carolyn fixed Yolanda up on a blind date. It was a very nice day. The weather was perfect for the first date. Enjoying being a match-maker, she looked forward to the date.
31. Elizabeth picked up the washing for Stephanie before the first guests arrived. The house was very dirty. It has not been cleaned for almost a month. Glad to do the favour, she thought about the special friendship.

32. Gina fixed a martini for Anna after watching a romantic movie. The movie was made by a new director. It became very popular. Pretending to be a bartender, she playfully stabbed a cocktail olive.
33. Pamela picked Stacey some yellow and white flowers. The flowers were all roses. They smelt very nice. After gathering a bouquet, she arranged them in a vase.
34. Kyle voted for Erik in the last presidential election. There were ten presidential candidates. The election was held in January. After winning the election, he was excited about the future.
35. Sara threw a piece of cake at Erin during the university graduation ceremony. The ceremony was being held at a stadium. It was a huge event. Not being a good target, she watched the cake hit the wall.
36. Amanda invited Andrea to a nice university party. It was for the start of the new academic year. It was supposed to be held at the university theatre. After accepting the invitation, she hoped it would be fun.
37. Kevin built Scott a nice wooden dog house. The dog house was very big. Five dogs could easily sleep in it. After finishing the dog house, he hoped the dogs liked it.
38. Suzanne laughed very loudly at Vanessa in a very important meeting. Lots of people were present at the meeting. It was an important meeting. Out of breath from laughing, she went quiet for a minute.
39. Shane passed the ball to Larry in a difficult football game. The game was against one of the best teams. That made it very hard. After running for a try, he envisioned the possibility of victory.
40. Dana received a chain letter from Jill with stories about sick children. The message was originally from a charity organization. The organization was raising money. After sending the letter, she practically forgot all about it.
41. Samantha reminded Courtney to do the dirty dishes. It was very late at night. But there were lots of dirty dishes in the kitchen. Hating receiving such reminders, she got in a bad mood.
42. Sharon made a cake for Brenda using a good new recipe. It was a big berry cheesecake. The cake smelt really good. After baking for two hours, she hoped the cheesecake was delicious.
43. Craig lost some money to Peter in gambling at the casino. The casino was located in a big hotel. It was always very busy. Enjoying the victory, he started walking toward the restaurant.
44. Charles saw Anthony have an accident while driving. The street was very narrow. There had always been lots of accidents on that street. Calling out in pain, he needed to find some help.
45. Alexander convinced Christian to apply for a job. A new library had just been built. Some staff was to be hired for the new library. After spending several hours convincing, he waited to hear the decision.
46. Johnny watched Philip swimming in the outdoor pool. It was a very hot day. Swimming was a good way to stay cool. After swimming for several hours, he got a drink of water.
47. Tammy tempted Laura with a piece of cake. The cake was very moist. It looked really delicious. Giving in to the temptation, she thought about all the calories.
48. Jennifer made Michelle a Chinese dish for lunch. The dish was very popular in China. It was mainly made of fish and rice. Using an old fashioned recipe, she knew it would taste good.

APPENDIX C

Sentence Stimuli for Comprehension task: Experiment 5

Experimental Paragraphs

1. Pamela picked Stacey some yellow and white flowers. The flowers were all roses. They smelt very nice. She gathered a very big bouquet.
2. Curtis got a postcard from Victor with a picture of beach. The sandy beach in the picture was full of shells. That made it more beautiful. He was jealous about the vacation.
3. Rachel drew a picture of Carrie sitting in a beautiful park. The scenery was amazing. The park was full of trees and flowers. She patiently posed for several hours.
4. Courtney reminded Samantha to do the dirty dishes. It was very late at night. But there were lots of dirty dishes in the kitchen. She hated to receive such reminders
5. Elizabeth picked up the washing for Stephanie before the first guests arrived. The house was very dirty. It has not been cleaned for almost a month. She really enjoyed doing the favour.
6. Gina fixed a martini for Anna after watching a romantic movie. The movie was made by a new director. It became very popular. She pretended to be a bartender.
7. Lee lost a book that belonged to Roy in the park last night. It was an expensive academic book. It was very old and really rare. He apologized for losing the book.
8. Brandy called the firemen to save Alicia from a burning university building. The fire was started by an oven explosion. It spread very quickly to all buildings. She was waiting to be rescued.
9. Cheryl helped Leslie across the very wide stream. It was a very beautiful stream. However, it was very deep. She was glad to be assisted.
10. Vanessa laughed very loudly at Suzanne in a very important meeting. Lots of people were present at the meeting. It was an important meeting. She felt ashamed of laughing loudly.
11. Jennifer made Michelle a Chinese dish for lunch. The dish was very popular in China. It was mainly made of fish and rice. She used an old fashioned recipe.
12. Dawn greeted Tina with smiles at the airport. The airport was very busy. A number of flights had landed. She was surprised by the welcome.
13. Larry passed the ball to Shane in a difficult football game. The game was against one of the best teams. That made it very hard. He lost the ball very quickly.
14. Kevin built Scott a nice wooden dog house. The dog house was very big. Five dogs could easily sleep in it. He finished the dog house quickly.
15. Rodney borrowed some money from Carlos to buy an expensive watch. It was a beautiful golden watch. It was on sale for a limited time. He was grateful for the loan.
16. Brian got a letter from James along with twenty five photos. The photos were very beautiful. They pictured the beautiful scenery of Mexico. He really enjoyed vacationing in Mexico.
17. Keith taught Bryan how to sing a song. It was a very beautiful love song. The song was very famous among the teenagers. He was a very good learner.

18. Brenda made a cake for Sharon using a new healthy recipe. It was a big berry cheesecake. The cake smelt really good. She was a very good baker.
19. Johnny watched Philip swimming in the outdoor pool. It was a very hot day. Swimming was a good way to stay cool. He swam for almost two hours.
20. Craig lost some money to Peter when gambling at the casino. The casino was located in a big hotel. It was always very busy. He enjoyed the victory very much.
21. Monica mailed a letter to Sandra using the international express service. The letter was a job offer. It contained some information about the job and the salary. She sent the letter first class.
22. Carolyn fixed Yolanda up on a blind date. It was a very nice day. The weather was perfect for the first date. She enjoyed being a match-maker.
23. Luis stole the basketball from Alan in a very competitive game. There were some valuable prizes for the winning teams. The best prize was a trip to Sydney. He cried after losing the ball.
24. Shawn handed the telephone to Aaron to call the hotel reception. It was the best hotel in the city. Therefore it was always fully booked. He picked up the receiver quickly.

APPENDIX D

Sentence Stimuli for Comprehension task: Experiment 6

Experimental Paragraphs

1. Lee greeted Roy with kisses, hugs and smiles. While giving the warm welcome, he began to get teary eyed.
2. Brenda nominated Sharon for class president this year. After making the nomination, she was excited about the future.
3. Cheryl watched Leslie jog around the big park. After jogging several laps, she got a drink of water.
4. Shawn made a big cake for Aaron to take to the party. Receiving the big cake, he knew it would taste good.
5. Monica loaned some money to Sandra to pay the electricity bill. Generous with the loan, she felt a sense of comradely.
6. Bryan told Keith a secret about a friend. After swearing not to tell anyone, he kept the secret strictly confidential.
7. Philip read an article for Johnny about nutrition and healthy diet. Carefully listening to the article, he wanted to follow the tips.
8. Yolanda handed the telephone to Carolyn to talk to a friend. After letting go of the receiver, she sat down on a chair.
9. Peter yelled at Craig for leaving the door unlocked. Not enjoying being yelled at, he was sorry the incident occurred.
10. Larry saw Shane fall down from a bike. Calling out in pain, he needed to find some help.
11. Courtney convinced Samantha to apply to medicine school. After spending hours on the application, she waited to hear the decision.
12. Rachel locked Carrie out of the house yesterday. After breaking in through a window, she put a key under the mat.
13. Tina dunked Dawn in the deep swimming pool. After coming up from the water, she reached for the pool side.
14. James tutored Brian in clinical and behavioural psychology. Never having been very good in psychology, he really enjoyed the tutoring session.
15. Curtis threw a big cream pie at Victor after the dinner last night. Not having a good aim, he watched the pie hit the wall.
16. Brandy rescued Alicia from a burning wooden building. Eternally grateful, she talked about it for years.
17. Gina cleaned the house for Anna after the surprise birthday party. After finishing the housework, she took a long afternoon nap.
18. Rodney loaned twenty dollars to Carlos to pay for the accommodation. Able to spare the cash, he felt good about the transaction.
19. Vanessa repeated the question for Suzanne at the oral history exam. Not having spoken clearly the first time, she tried even harder to concentrate.

20. Stephanie lost a tennis match to Elizabeth in the big old stadium. Accepting the defeat, she started walking toward the showers.
21. Michelle taught Jennifer how to draw a horse. Being a good teacher, she made the job seem easy.
22. Alan told Luis the awful truth last week. After having heard it, he hoped it wouldn't be repeated.
23. Pamela reminded Stacey to take out the garbage. Hating having to give such reminders, she got in a bad mood.
24. Kevin told Scott about a very funny cartoon After giving the review, he suggested watching the funny cartoon.

REFERENCES

Allen, H., Ledgeway, T., Kelly, N., Hutchinson, C., & Blundell, J. (2011). Divided attention impairs motion perception in older adults. *Journal of Vision, 11*(11), 97. doi: 10.1167/11.11.97

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: a proposed system and its control processes. In W. S. Kenneth & S. Janet Taylor (Eds.), *Psychology of Learning and Motivation* (Vol. Volume 2, pp. 89-195): Academic Press.

Au, R., Albert, M. L., & Obler, L. K. (1989). Language in normal aging: linguistic and neuropsychological factors. *Journal of Neurolinguistics, 4*(3–4), 347-364.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412. doi: http://dx.doi.org/10.1016/j.jml.2007.12.005

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *Análisis de los tiempos de reacción., 3*(2), 12-28.

Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders, 36*(3), 189-208.

Baddeley, A. (2010). Working memory. [doi: 10.1016/j.cub.2009.12.014]. *Current Biology, 20*(4), R136-R140.

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology, 63*(1), 1-29. doi: doi:10.1146/annurev-psych-120710-100422

Baddeley, A., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). New York: Academic Press.

Bell, R., Buchner, A., & Mund, I. (2008). Age-related differences in irrelevant-speech effects. *Psychology and Aging, 23*(2), 377-391.

Bergland, A., & Narum, I. (2007). Quality of life demands comprehension and further exploration. *Journal of Aging and Health, 19*(1), 39-61. doi: 10.1177/0898264306296766

Blasi, G., Goldberg, T. E., Weickert, T., Das, S., Kohn, P., Zoltick, B., . . . Mattay, V. S. (2006). Brain regions underlying response inhibition and interference monitoring and suppression. [Article]. *European Journal of Neuroscience, 23*(6), 1658-1664. doi: 10.1111/j.1460-9568.2006.04680.x

Borella, E., Ghisletta, P., & De Ribaupierre, A. (2011). Age differences in text processing: the role of working memory, inhibition, and processing speed. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences, 66 B*(3), 311-320.

Brébion, G. (2003). Working memory, language comprehension, and aging: four experiments to understand the deficit. [doi: 10.1080/03610730303725]. *Experimental Aging Research, 29*(3), 269-301. doi: 10.1080/03610730303725

Brydges, C. R., Clunies-Ross, K., Clohessy, M., Lo, Z. L., Nguyen, A., Rousset, C., . . . Fox, A. M. (2012). Dissociable components of cognitive control: an Event-Related Potential (ERP) study of response inhibition and interference suppression. *PLoS ONE, 7*(3), e34482. doi: 10.1371/journal.pone.0034482

Burke, D. M., & Shafto, M. A. (2008). Language and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The Handbook of Aging and Cognition* (third ed.). New York: Psychology Press.

Caplan, D. (1972). Clause boundaries and recognition latencies for words in sentences. *Perception & psychophysics, 12*(1), 73 - 76.

Carpenter, P. A., Miyake, A., & Just, M. A. (1995). Language comprehension: sentence and discourse processing. *Annual Review of Psychology, 46*, 91-91.

Carreiras, M., Gernsbacher, M. A., & Villa, V. (1995). The advantage of first mention in Spanish. *Psychonomic Bulletin & Review, 2*(1), 124-129. doi: 10.3758/BF03214418

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences, 19*(2), 246-251. doi: http://dx.doi.org/10.1016/j.lindif.2008.10.002

Carretti, B., Cornoldi, C., De Beni, R., & Romanò, M. (2005). Updating in working memory: a comparison of good and poor comprehenders. *Journal of Experimental Child Psychology, 91*(1), 45-66. doi: http://dx.doi.org/10.1016/j.jecp.2005.01.005

Carroll, D. W. (2008). *Psychology of language* (5th ed.). Belmont, CA: ThomsonWadsworth.

Cohen, G. (1979). Language comprehension in old age. *Cognitive Psychology, 11*(4), 412-429.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of active maintenance and executive control* (pp. 62-101): Cambridge University Press.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450-466.

De Beni, R., Borella, E., & Carretti, B. (2007). Reading comprehension in aging: the role of working memory and metacomprehension. [doi: 10.1080/13825580500229213]. *Aging, Neuropsychology, and Cognition, 14*(2), 189-212. doi: 10.1080/13825580500229213

Duffy, S. A., & Keir, J. A. (2004). Violating stereotypes: eye movements and comprehension processes when text conflicts with world knowledge. *Memory & Cognition, 32*(4), 551-559. doi: 10.3758/bf03195846

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: a test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(5), 972-992. doi: 10.1037/0278-7393.18.5.972

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology: General, 128*(3), 309-331.

Federmeier, K. D., & Kutas, M. (2005). Aging in context: age-related changes in context use during language comprehension. *Psychophysiology, 42*(2), 133-141. doi: 10.1111/j.1469-8986.2005.00274.x

Federmeier, K. D., van Petten, C., Schwartz, T. J., & Kutas, M. (2003). Sounds, words, sentences: age-related changes across levels of language processing. *Psychology and Aging, 18*(4), 858-872.

Finnigan, S., O'Connell, R. G., Cummins, T. D. R., Broughton, M., & Robertson, I. H. (2011). ERP measures indicate both attention and working memory encoding decrements in aging. *Psychophysiology, 48*(5), 601-611.

Fodor, J. (1983). *The modularity of mind*. Cambridge , MA: MIT Press.

Friedman, N. P., & Miyake, A. (2004a). The reading span test and its predictive power for reading comprehension ability. *Journal of Memory and Language, 51*(1), 136-158. doi: http://dx.doi.org/10.1016/j.jml.2004.03.008

Friedman, N. P., & Miyake, A. (2004b). The relations among inhibition and interference control functions: a latent-variable analysis. *Journal of Experimental Psychology: General, 133*(1), 101-135. doi: 10.1037/0096-3445.133.1.101

Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition, 32*(2), 99-156.

Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Gernsbacher, M. A. (1991). Cognitive processes and mechanisms in language comprehension: the structure building framework. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 27, pp. 217-263). New York: Academic Press.

Gernsbacher, M. A. (1997a). Attentuating interference during comprehension: the role of suppression *Psychology of Learning and Motivation - Advances in Research and Theory* (Vol. 37, pp. 85-104).

Gernsbacher, M. A. (1997b). Group differences in suppression skill. *Aging, Neuropsychology, and Cognition, 4*(3), 175-184. doi: 10.1080/13825589708256646

Gernsbacher, M. A. (1997c). Two decades of structure building. *Discourse Processes, 23*(3), 265-304. doi: 10.1080/01638539709544994

Gernsbacher, M. A., & Faust, M. E. (1991). The mechanism of suppression: a component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(2), 245-262. doi: 10.1037/0278-7393.17.2.245

Gernsbacher, M. A., & Foertsch, J. (1999). Three models of discourse comprehension. In S. G. M. Pickering (Ed.), *Language processing* (pp. 283-299). East Sussex, UK: Psychology Press.

Gernsbacher, M. A., & Hargreaves, D. J. (1988). Accessing sentence participants: the advantage of first mention. *Journal of Memory and Language, 27*(6), 699-717. doi: http://dx.doi.org/10.1016/0749-596X(88)90016-2

Gernsbacher, M. A., Hargreaves, D. J., & Beeman, M. (1989). Building and accessing clausal representations: the advantage of first mention versus the advantage of clause recency. *Journal of Memory and Language, 28*(6), 735-755. doi: http://dx.doi.org/10.1016/0749-596X(89)90006-5

Gernsbacher, M. A., & Jescheniak, J. D. (1995). Cataphoric devices in spoken discourse. *Cognitive Psychology, 29*(1), 24-58.

Gernsbacher, M. A., & Kaschak, M. (2006). Discourse processing. In B. Keith (Ed.), *Encyclopedia of Language & Linguistics* (pp. 654-660). Oxford: Elsevier.

Gernsbacher, M. A., Keysar, B., Robertson, R. R. W., & Werner, N. K. (2001). The role of suppression and enhancement in understanding metaphors. *Journal of Memory and Language, 45*(3), 433-450.

Gernsbacher, M. A., & Robertson, R. R. W. (1999). The role of suppression in figurative language comprehension. *Journal of Pragmatics, 31*(12), 1619-1630. doi: http://dx.doi.org/10.1016/S0378-2166(99)00007-7

Gernsbacher, M. A., Tallent, K. A., & Bolliger, C. M. (1999). Disordered discourse in schizophrenia described by the structure building framework. *Discourse Studies, 1*(3), 355-372.

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(3), 430-445. doi: 10.1037/0278-7393.16.3.430

Giaquinto, S., Ranghi, F., & Butler, S. (2007). Stability of word comprehension with age: an electrophysiological study. *Mechanisms of Ageing and Development, 128*(11–12), 628-636.

Giora, R. (1996). Language comprehension as structure building. *Journal of Pragmatics, 26*(3), 417-436.

Goral, M. (2004). First-language decline in healthy aging: implications for attrition in bilingualism. *Journal of Neurolinguistics, 17*(1), 31-52.

Gordon, P. C., Hendrick, R., & Foster, K. L. (2000). Language comprehension and probe-list memory. *Journal of Experimental Psychology: Learning Memory and Cognition, 26*(3), 766-775.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48*, 163-189.

Greene, S. B., McKoon, G., & Ratcliff, R. (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(2), 266-283.

Hamm, V. P., & Hasher, L. (1992). Age and the availability of inferences. *Psychology and Aging, 7*(1), 56-64. doi: 10.1037/0882-7974.7.1.56

Hancock, H. E., Fisk, A. D., & Rogers, W. A. (2001). Comprehension of explicit and implicit warning information in younger and older adults. *Proceedings of the Human Factors and Ergonomics Society ... Annual Meeting, 2*, 1474-1474.

Hannon, B., & Daneman, M. (2009). Age-related changes in reading comprehension: an individual-differences perspective. *Experimental Aging Research, 35*(4), 432-456. doi: 10.1080/03610730903175808

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

Hermans, D., De Houwer, J., & Eelen, P. (2001). A time course analysis of the affective priming effect. *Cognition & Emotion, 15*(2), 143-165. doi: 10.1080/02699930125768

Hermans, D., Spruyt, A., & Eelen, P. (2003). Automatic affective priming of recently acquired stimulus valence: priming at SOA 300 but not at SOA 1000. *Cognition & Emotion, 17*(1), 83-99. doi: 10.1080/02699930302276

Järvikivi, J., van Gompel, R. P. G., Hyönä, J., & Bertram, R. (2005). Ambiguous pronoun resolution: contrasting the first-mention and subject-preference accounts. *Psychological Science, 16*(4), 260-264. doi: 10.1111/j.0956-7976.2005.01525.x

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review, 99*(1), 122-149. doi: 10.1037/0033-295X.99.1.122

Kensinger, E. A. (2009). Cognition in aging and age-related disease. In P. R. Hof, Mobbs, C. V. (Ed.), *Handbook of the Neuroscience of Aging*. London, UK: Elsevier Press.

Kim, S.-i., Lee, J.-h., & Gernsbacher, M. A. (2004). The advantage of first mention in korean: the temporal contributions of syntactic, semantic, and pragmatic factors. *Journal of Psycholinguistic Research, 33*(6), 475-491. doi: 10.1007/s10936-004-2667-5

Krawietz, S. A., Tamplin, A. K., & Radvansky, G. A. (2012). Aging and mind wandering during text comprehension. *Psychology and Aging*. doi: 10.1037/a0028831

Kynette, D., & Kemper, S. (1986). Aging and the loss of grammatical forms: a cross-sectional study of language performance. *Language & Communication, 6*(1–2), 65-72.

Lee, C. L., & Federmeier, K. D. (2012). Ambiguity's aftermath: how age differences in resolving lexical ambiguity affect subsequent comprehension. *Neuropsychologia, 50*(5), 869-879. doi: http://dx.doi.org/10.1016/j.neuropsychologia.2012.01.027

Lee, D., & Newman, S. D. (2010). The effect of presentation paradigm on syntactic processing: an event-related fMRI study. *Hum Brain Mapp, 31*(1), 65-79. doi: 10.1002/hbm.20845

Light, L. L., & Capps, J. L. (1986). Comprehension of pronouns in young and older adults. *Developmental Psychology, 22*(4), 580-585.

Love, J., & McKoon, G. (2011). Rules of Engagement: incomplete and Complete Pronoun Resolution. *Journal of Experimental Psychology: Learning Memory and Cognition, 37*(4), 874-887.

Lustig, C., Hasher, L., & Zacks, R. T. (2007). Inhibitory deficit theory: recent developments in a "new view". In D. S. Gorfein & C. M. MacLeod (Eds.), *Inhibition in cognition* (Vol. xvii, pp. 145-162). Washington, DC, US: American Psychological Association.

Malmstrom, T., & LaVoie, D. J. (2002). Age Differences in Inhibition of Schema-Activated Distractors. *Experimental Aging Research, 28*(3), 281-298. doi: 10.1080/03610730290080335

Maxim, J., & Bryan, K. (1994). *Language of the elderly*. London: Whurr.

McDonald, J. L., & Shaibe, D. M. (2002). The accessibility of characters in single sentences: proper names, common nouns, and first mention. *Psychonomic Bulletin and Review, 9*(2), 356-361.

Milton, D. J. (2008). *Working memory and academic learning : assessment and intervention* Hoboken, N.J. : John Wiley & Sons, Inc.

Morrone, I., Declercq, C., Novella, J. L., & Besche, C. (2010). Aging and inhibition processes: the case of metaphor treatment. *Psychology and Aging, 25*(3), 697-701.

Newman, S. D., Malaia, E., Seo, R., & Cheng, H. (2013). The effect of individual differences in working memory capacity on sentence comprehension: an FMRI study. *Brain Topography, 26*(3), 458-467.

Nicholas, M., Obler, L., Albert, M. L., & Goodglass, H. (1985). Lexical retrieval in healthyaging. *Cortex, 21*, 595-606.

Noh, S., & Stine-Morrow, E. A. L. (2009). Age differences in tracking characters during narrative comprehension. *Memory & Cognition, 37*(6), 769-778. doi: 10.3758/mc.37.6.769

Norman, S., Kemper, S., & Kynette, D. (1992). Adults' reading comprehension: effects of syntactic complexity and working memory. *Journal of Gerontology, 47*(4), 8.

Obler, L. K., Fein, D., Nicholas, M., & Albert, M. L. (1991). Auditory comprehension and aging: decline in syntactic processing. *Applied Psycholinguistics, 12*(04), 433-452. doi: doi:10.1017/S0142716400005865

Park, D. C., Denise, C. P., Anderson, D. S., William, N. D., & Vincent, N. L. (1989). Effects of age and a divided attention task presented during encoding and retrieval on memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(6), 1185-1191. doi: 10.1037/0278-7393.15.6.1185

Perfetti, C. A. (1994). Psycholinguistics and reading ability. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 849-894). San Diego: Academic Press.

Prat, C. S., Keller, T. A., & Just, M. A. (2007). Individual differences in sentence comprehension: a functional magnetic resonance imaging investigation of syntactic and lexical processing demands. *Journal of Cognitive Neuroscience, 19*(12), 1950-1963. doi: 10.1162/jocn.2007.19.12.1950

Radvansky, G. A. (1999). Aging, memory, and comprehension. *Current Directions in Psychological Science, 8file:///R:/fmri/20182557.pdf*(2), 49-53. doi: 10.2307/20182557

Radvansky, G. A., Copeland, D. E., Berish, D. E., & Dijkstra, K. (2003). Aging and situation model updating. [doi: 10.1076/anec.10.2.158.14459]. *Aging, Neuropsychology, and Cognition, 10*(2), 158-166. doi: 10.1076/anec.10.2.158.14459

Radvansky, G. A., Copeland, D. E., & Hippel, W. v. (2010). Stereotype activation, inhibition, and aging. *Journal of Experimental Social Psychology, 46*(1), 51-60.

Radvansky, G. A., & Dijkstra, K. (2007). Aging and situation model processing. *Psychonomic Bulletin & Review, 14*(6), 1027-1042. doi: 10.3758/bf03193088

Radvansky, G. A., Lynchard, N., & von Hippel, W. (2009). Aging and stereotype suppression. [doi:10.1080/13825580802187200]. *Aging, Neuropsychology, and Cognition (Neuropsychology, Development and Cogniti, 16*(1), 22-32.

Radvansky, G. A., Zwaan, R. A., Curiel, J. M., & Copeland, D. E. (2001). Situation models and aging. *Psychology and Aging, 16*(1), 145-160. doi: 10.1037/0882-7974.16.1.145

Rayner, K., & Clifton, C. J. (2002). Language processing. In H. Pashler & D. Medin (Eds.), *Stevens Handbook of Experimental Psychology: Memory and Cognitive Processes* (third ed., Vol. 2, pp. 261-316). New York: John Wiley and Sons, Inc.

Salthouse, T. A. (1992). Why do adult age differences increase with task complexity? *Developmental Psychology, 28*(5), 905-918. doi: 10.1037/0012-1649.28.5.905

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*(3), 403-428.

Shake, M. C., & Stine-Morrow, E. A. L. (2011). Age differences in resolving anaphoric expressions during reading. *Aging, Neuropsychology, and Cognition, 18*(6), 678-707. doi: 10.1080/13825585.2011.607228

Shewan, C. M., & Henderson, V. L. (1988). Analysis of spontaneous language in the older normal population. *Journal of Communication Disorders, 21*(2), 139-154.

Social Security. (2013). *Top names of the 1970s.* Retrieved from http://www.ssa.gov/oact/babynames/decades/names1970s.html

Stine-Morrow, E. A. L., Ryan, S., & Leonard, J. S. (2000). Age differences in on-line syntactic processing. *Experimental Aging Research, 26*(4), 315-322. doi: 10.1080/036107300750015714

Streb, J., Hennighausen, E., & Rösler, F. (2004). Different anaphoric expressions are investigated by event-related brain potentials. *Journal of Psycholinguistic Research, 33*(3), 175-201.

Susan, A. D., & David, B. P. (1992). Comprehension of synthetic speech produced by rule: A review and theoretical interpretation. [Article]. *Language & Speech, 35*(4), 351-389.

Towse, J. N., Hitch, G. J., & Hutton, U. (2000). On the interpretation of working memory span in adults. *Memory & Cognition, 28*(3), 341-348. doi: 10.3758/BF03198549

Uekermann, J., Thoma, P., & Daum, I. (2008). Proverb interpretation changes in aging. *Brain and Cognition, 67*(1), 51-57. doi: http://dx.doi.org/10.1016/j.bandc.2007.11.003

Ulatowska, H. K., Hayashi, M. M., Cannito, M. P., & Fleming, S. G. (1986). Disruption of reference in aging. *Brain and Language, 28*(1), 24-41.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*, 498-505.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

van Rij, J., van Rijn, H., & Hendriks, P. (2013). How WM load influences linguistic processing in adults: a computational model of pronoun interpretation in discourse. *Topics in Cognitive Science, 5*(3), 564-580.

Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology Section A, 47*(3), 631-650. doi: 10.1080/14640749408401131

von Hippel, W., Silver, L. A., & Lynch, M. E. (2000). Stereotyping against your will: the role of inhibitory ability in stereotyping and prejudice among the elderly. *Personality and Social Psychology Bulletin, 26*(5), 523-532. doi: 10.1177/0146167200267001

Was, C. A., Rawson, K. A., Bailey, H., & Dunlosky, J. (2011). Content-embedded tasks beat complex span for predicting comprehension. *Behavior Research Methods, 43*(4), 910-915. doi: 10.3758/s13428-011-0112-x

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review, 72*(2), 89-104.

Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record, 58*(3), 475-482. doi: 10.1037/0096-3445.128.1.32_10.1037/0096-3445.123.1.34

Wingfield, A., Peelle, J. E., & Grossman, M. (2003). Speech rate and syntactic complexity as multiplicative factors in speech comprehension by young and older adults. *Aging, Neuropsychology, and Cognition, 10*(4), 310-322. doi: 10.1076/anec.10.4.310.28974