THE EFFECTS OF AGE AND VISUAL STIMULI ON NARRATIVE PRODUCTION

by

CLINTON YIN HANG TSANG
B.A., University of British Columbia, 2007

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES
(Audiology and Speech Sciences)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)
December, 2009

© Clinton Yin Hang Tsang, 2009
Abstract

Narrative production has been used as a tool for assessing and treating individuals with acquired language disorders as well as cognitive-linguistic deficits. Because individuals who sustained these conditions tend to be older adults, it is important to distinguish between the effects of normal aging versus the effects of these conditions on narrative production. Research has also suggested that both aging and the use of different visual materials may influence performance on narrative production tasks. This study was designed to investigate the effects of aging and the use of different visual stimuli, namely, picture sequence versus video, in narrative production by normal healthy adults.

Twenty younger and twenty older adults with no cognitive or language deficits participated in this study. Each participant produced two different narratives, one based on a picture sequence stimulus and the other based on a video stimulus. Lexical diversity, verb-noun ratio, cohesion, coherence and content richness of each narrative were measured and compared across age groups and task conditions. Results of the study showed that age did not have an effect on any of the dependent measures. Narratives elicited from the picture sequence condition appeared to have a higher lexical diversity than those elicited from the video condition, which was found to be due to the lower total number of words in narratives of the picture sequence condition. Task did not have an effect on verb-noun ratio, cohesion, coherence and content richness. Unexpectedly, a main effect of task was found on total number of words and the percentage of core propositions produced. Limitations and clinical implication of the current study, as well as future direction for research, are discussed.
Table of Contents

Abstract ................................................................................................................................. ii

Table of Contents ................................................................................................................ iii

List of Tables ........................................................................................................................ viii

List of Figures ....................................................................................................................... ix

Acknowledgements .............................................................................................................. x

1. Introduction ......................................................................................................................... 1

1.1 Narrative as a Genre of Discourse ................................................................................. 2

1.1.1 Elicitation and structure of narrative ....................................................................... 2

1.1.2 Cognitive skills related to narrative production ..................................................... 4

1.1.2.1 Working memory ............................................................................................... 5

1.1.2.2 Inhibition ........................................................................................................... 6

1.1.2.3 Inference ............................................................................................................ 6

1.1.2.4 Symbol processing ............................................................................................ 8

1.2 Narrative in Clinical Assessment ...................................................................................... 9

1.2.1 Clinical assessment and research using narrative production ................................... 9

1.2.2 Analysing narratives ................................................................................................. 10

1.2.2.1 Lexical measures .............................................................................................. 11

1.2.2.2 Cohesion and coherence ................................................................................. 12

1.2.2.3 Content richness ............................................................................................. 13

1.3 Age-related Differences in Narrative Production ......................................................... 14

1.3.1 Age and lexical processing ......................................................................................... 15

1.3.1.1 Lexical diversity and access ....................................................................... 15
1.3.1.2 Verb-noun ratio ................................................................................. 18
1.3.2 Age and cohesion and coherence ......................................................... 20
  1.3.2.1 Cohesion ......................................................................................... 21
  1.3.2.2 Coherence ....................................................................................... 23
1.3.3 Age and content richness ...................................................................... 25
1.4 The Use of Different Visual Stimuli in Discourse Elicitation ..................... 28
  1.4.1 Potential differences among single picture, picture sequence and video .... 29
  1.4.2 Single picture vs. picture sequence ...................................................... 31
    1.4.2.1 Cohesion ......................................................................................... 31
    1.4.2.2 Coherence ....................................................................................... 31
    1.4.2.3 Content richness .............................................................................. 32
  1.4.3 Picture sequence vs. video ................................................................. 33
    1.4.3.1 Visual stimuli and narrative production in children ......................... 34
    1.4.3.2 Pictures and videos in language acquisition ..................................... 35
1.5 Summary of Literature Review ................................................................. 36
1.6 Research Hypotheses ................................................................................. 37
2. Methods ........................................................................................................ 40
  2.1 Participants .............................................................................................. 40
    2.1.1 Demographic information ................................................................. 40
    2.1.2 Working memory span ..................................................................... 41
      2.1.2.1 Sentence span task ..................................................................... 42
      2.1.2.2 Counting span task ..................................................................... 44
    2.1.3 Lexical knowledge ............................................................................. 46
3.5 Summary of Findings ........................................................................................................ 76

4. Discussion .......................................................................................................................... 78

4.1 Research Hypotheses ........................................................................................................ 78

4.1.1 The effects of age on lexical diversity, verb-noun ratio, and coherence ................. 78

4.1.1.1 Age and lexical diversity .................................................................................. 79

4.1.1.2 Age and verb-noun ratio ................................................................................ 80

4.1.1.3 Age and coherence ......................................................................................... 80

4.1.2 The effects of age on cohesion and content richness .............................................. 81

4.1.2.1 Age and cohesion ......................................................................................... 81

4.1.2.2 Age and content richness ............................................................................. 82

4.1.3 The effects of task on cohesion ............................................................................. 83

4.1.4 The effects of task on coherence and content richness ............................................ 84

4.2 Other Research Findings: Insights into Dependent Variable Measures ..................... 85

4.2.1 Type-token ratio: the effect of total number of words .......................................... 86

4.2.1.1 Number of different words ........................................................................... 86

4.2.1.2 D value ........................................................................................................ 87

4.2.2 Total number of words as a primary dependent variable ....................................... 88

4.2.3 Floor effect of cohesion and coherence error ......................................................... 88

4.2.4 Core proposition analysis of the Cowboy’s Story narratives ................................ 90

4.3 Limitations ...................................................................................................................... 92

4.4 Clinical Implications ...................................................................................................... 94

4.5 Directions for Future Research ..................................................................................... 95

4.6 Conclusion ....................................................................................................................... 96
References .................................................................................................................. 98
APPENDIX A: Ethics certificate ..................................................................................... 107
APPENDIX B: Participant information and informed consent ................................. 109
APPENDIX C: NAART35 word list .............................................................................. 112
APPENDIX D: Fisherman’s Tale .................................................................................... 113
APPENDIX E: Sample transcript ................................................................................... 114
APPENDIX F: Definitions and CLAN codes for cohesion and coherence errors .......... 115
APPENDIX G: Sample proposition analysis of transcript in Appendix E ................... 116
APPENDIX H: Core propositions of the Cowboy’s Story from Joanette and Goulet (1990).... 118
List of Tables

Table 2.1 Demographic information about participants .........................................................41
Table 2.2 Working memory span scores .............................................................................46
Table 2.3 NAART35 raw and z-score ................................................................................48
Table 3.1 Summary of data arranged by story, task condition and age group ......................59
Table 3.2 Summary of data arranged by age group and task condition ...............................60
Table 3.3 Marginally significant effects of order and combination .......................................67
Table 3.4 Comparison of the collective vs. the separate analyses of the two stories .............74
List of Figures

Figure 3.1 TTR by age ........................................................................................................... 61
Figure 3.2 TTR by task ......................................................................................................... 61
Figure 3.3 TNW by task ....................................................................................................... 62
Figure 3.4 Verb-noun ratio by age ....................................................................................... 63
Figure 3.5 Verb-noun ratio by task ...................................................................................... 63
Figure 3.6 Cohesion error by age ......................................................................................... 64
Figure 3.7 Cohesion error by task ....................................................................................... 64
Figure 3.8 Coherence error by age ....................................................................................... 65
Figure 3.9 Coherence error by task ...................................................................................... 65
Figure 3.10 P-density by age ............................................................................................... 66
Figure 3.11 P-density by task .............................................................................................. 66
Figure 3.12 Cowboy’s Story TTR by task .............................................................................. 69
Figure 3.13 Cowboy’s Story cohesion error by task .............................................................. 70
Figure 3.14 Fisherman’s Tale TTR by task ......................................................................... 72
Figure 3.15 Fisherman’s Tale verb-noun ratio by age and task ............................................. 73
Figure 3.16 Mean percentage of core propositions produced by age .................................. 75
Figure 3.17 Mean percentage of core propositions produced by task ............................... 75
Acknowledgements

To Dr. Barbara A. Purves, thank you so much for taking me as your thesis student and being such an enthusiastic, inspiring and supportive supervisor. Your passion for research and the way you truly care about your students and their work have left me completely speechless and in awe. Your gentle and reassuring encouragements have also helped me survive my numerous moments of insecurity throughout the whole process. I have learnt so much from you in the last two years and words are just not enough to express how grateful I am to have you as my supervisor.

To Dr. Jeff A. Small, thank you for sitting on my committee and for all the invaluable advice you have given me regarding the study design and statistical analysis. Thank you also for dropping by the lab from time to time just to see how I was doing, as it reassured me that someone does care about my research as well as my well-being.

To Dr. Stefka Marinova-Todd, thank you for sitting on my committee and providing guidance in the process of statistical analysis. Thank you also for allowing me to work in the Bilingualism Research Lab in the last two years where I have acquired precious experience in discourse research prior to the commencement of this study.

To Dr. Paolo Colozzo, thank you for helping me with the statistical analysis of my data and the interpretation of the results despite you busy schedule and heavy workload.

To my parents, thank you for providing me with a home and an environment that motivates me to always try my best and excel in whatever I do.
To my brother, Dr. Jenkins Yin Ki Tsang, thank you for always being so supportive of whatever I do and the best brother one can ever ask for.

To St. Paul’s Co-educational College, Hong Kong, and the Department of Linguistics, UBC, thank you for giving me a great secondary and post-secondary education that prepared and allowed me to pursue a career in Speech-Language Pathology.

To Dr. Peter Graf and Ashleigh Chapman from the Department of Psychology, UBC; Jason Brown from the Department of Linguistics, UBC; Speech and Linguistics Student Association, UBC; Jane Hack from the Dunbar Patrol; Brian Nielsen from Kerrisdale Lawn Bowling Club; and Kerrisdale Community Centre, thank you for helping me with the recruitment of participants.

To Dr. Gloria Waters and Jennifer Michaud, Boston University, thank you for sharing your materials of the sentence span task with me.

To Eri Yamamoto, thank you for preparing the picture stimuli for my study.

To Jill Petersen and Amy Makaroff, thank you for doing the reliability checks on the transcriptions and the analyses.

To Mijke Rhemtulla, thank you for your invaluable assistance in the statistical analyses of the data.

Lastly, to the School of Audiology and Speech Sciences Class of 2009, thank you for being there for me in the last two years. I would not have been able to get through the program without all of you.
1. Introduction

Narrative production is an integral part of our everyday life. Whenever we are recalling an event from the past, sharing personal experiences, or telling fairy tales to children, we are essentially producing narratives. The study of narratives is an invaluable tool as it provides insight into both social and linguistic behaviours. In terms of linguistic behaviours, it allows for a comprehensive investigation of individuals’ language use across multiple linguistic domains. For example, a comprehensive analysis of a narrative may yield information about aspects of language use, such as lexical diversity, morphological accuracy, syntactic complexity, and pragmatic skills. Because such variety of information can be gathered, narrative production has been used in clinical assessments of linguistic and cognitive abilities of individuals with acquired brain injury. However, it is important to first understand how normal healthy adults produce narrative in order for clinicians to be able to compare impaired discourse behaviours to normal discourse behaviours. Furthermore, because individuals with acquired brain injury such as cerebrovascular accidents (CVA) tend to be older, it is critical for clinicians to have some knowledge about discourse behaviours of older adults so that they are able to distinguish between effects of normal aging vs. effects of the CVA on discourse production. Therefore results of narrative research on normal healthy adults may reveal discourse characteristics that may serve as a baseline for comparison with clinical populations (Mackenzie, 2000; Ulatowska, Hayashi, Cannito, & Fleming, 1986).

Despite these tremendous values of studying narrative production, research in this area has been relatively scarce. The number of studies using normal healthy adults as participants is especially limited. Within this body of research, one particular area of interest
relevant to the current study is the effects of age on narrative production. A second area of interest is the use of different visual stimuli in the elicitation of narratives. This has been studied in the context of clinical practice, since it is important for clinicians to be aware of how different elicitation methods may influence the resulting narratives. This study investigates the effect of age and the use of picture sequence vs. video as visual stimuli on verb use, cohesion and coherence, and content richness in narrative production by normal healthy adults. In the following sections, the nature of narrative production and its place in clinical assessment are discussed. This is followed by a review of the current literature on the influence of age and visual stimuli on narrative production.

1.1 Narrative as a Genre of Discourse

Discourse can be defined as "naturally occurring language that extends beyond the sentence level or across sentences" (Ehrlich, 1994, p. 150). Discourse can be divided into different genres based on their macrostructure and/or the function they serve (Duchan, 1994). Some of these genres include conversation, narrative, persuasive talk, interview, courtroom discourse, and conflict resolution. Among these genres, narrative has been most commonly studied (Duchan, 1994).

1.1.1 Elicitation and structure of narrative

Narrative refers to an extended unit of oral or written language consisting of multiple propositions that collectively represent one or multiple events, as in stories, descriptions, explanations, and expositions (Ehrlich, 1994; Roth & Spekman, 1994). Unlike conversation, where there are multidirectional transactions of propositional information among the interlocutors on a topic that they implicitly or explicitly agree on, propositions are usually
delivered unidirectionally from the narrator(s) to the audience. Narrator(s) also possess some internal rule systems or frameworks that allow them to generate narrative of a particular structure, as well as strategies for adjusting to different listeners and audiences (Ehrlich, 1994).

In discourse research and clinical practice, narratives can be elicited using different tasks and procedures. For example, narratives can be elicited by a single picture description task (Duong & Ska, 2001; Liles, Coehlo, Duffy, & Zalagens, 1989; Marini, Boewe, Caltagirone, & Carlomagno, 2005; Wright, Capilouto, Wagovich, Cranfill, & Davis, 2005), where the individual is shown a picture or a scene and is asked to describe the event taking place in the stimulus. Alternatively, a picture sequence such as a comic strip or a wordless story book can be shown to an individual, who is then asked to produce a narrative that describes the event(s) taking place in the pictures (Duong & Ska, 2001; Marini et al. 2005; Wright et al., 2005). Story recall is also commonly used in narrative research (Liles et al., 1989; Pratt, Boyes, Robins, & Manchester, 1989). In this procedure, a story is presented visually in words or auditorily in speech to an individual. The individual is then asked to reiterate the story as closely to the original stimulus as possible. Finally, personal narratives can also be used, in which individuals are asked to recall or describe an event or a personal experience (Ulatowska et al., 1986).

While narratives can be elicited with different methods, the resulting narratives usually roughly consist of two components, namely, the microstructure and the macrostructure. Microstructure refers to the linguistic components of the narrative, such as vocabulary use and the use of different sentence structures. Macrostructure, on the other hand can be defined as the central theme and the overall organization of the propositional material.
of a given piece of discourse. In the case of narrative, researchers have proposed various schemes to characterize the components of the macrostructure (Labov, 1972; Stein & Glenn, 1979). In a review of such schemes, Ochs (1999) summarized the key constituents of the macrostructure of narratives as follows: (1) abstract, an overview statement of the narrative; (2) setting, which involves the specification of background information of the story such as context, environment, and characters; (3) complication, the development of the event(s) of the story; (4) evaluation, any statements of personal opinions on the event(s); (5) result or resolution, how the complications are resolved; and (6) coda, closing statement(s) of the story. A given narrative may have some or all of these elements, depending on factors such as the nature of the narrative and the narration skills of the narrator. For instance, setting, complication, and resolution tend to be more fundamental, and thus, present in most narratives, whereas abstract, evaluation, and coda tend to be optional.

1.1.2 Cognitive skills related to narrative production

Narrative production is an activity that demands both linguistic and cognitive processing. While linguistic processing allows an individual to build the microstructure of the narrative, such as using appropriate vocabulary and grammatical sentence structures, cognitive processing is crucial for comprehending and conceptualizing the visual stimulus (e.g., in the case of picture description tasks) and for constructing the macrostructure of the narrative. In this section, four of these cognitive skills, namely, working memory, inhibition, inference, and symbol processing, are discussed.
1.1.2.1 Working memory.

According to Baddeley and Hitch (1974), as cited by Waters and Caplan (2003), working memory is a cognitive system that is not used merely for the storage of a small amount of information in a short period of time (as in the case of short-term memory), but also carries out ongoing and simultaneous information processing and manipulation of the materials being remembered. Baddeley (1992) proposed a theoretical model to conceptualize the cognitive domain of working memory. In this model, working memory consists of a central executive that presumably controls attention-related processes and two slave systems. The visual-spatial slave system processes visual images, while the phonological loop manipulates auditory linguistic inputs. The central executive and the visual-spatial slave system mostly form the visual working memory while verbal working memory is largely made up of the central executive and the phonological loop slave system. Functional working memory capacity depends on the intactness of these three systems and their connections.

Evidence in the discourse literature suggests that working memory may be important to various aspects of narrative production, such as the maintenance of cohesion through the use of pronouns. For instance, Pratt et al. (1989) administered a working memory span task to the participants in their narrative study. They specifically found that working memory span size negatively correlates with the amount of pronoun error in narrative production. Such a finding is not surprising since the use of clear and accurate pronouns requires the speaker to keep track of the previously mentioned antecedents in the discourse, access the correct grammatical pronominal lexical item, and continue formulating the rest of the utterance. This places a huge demand on working memory capacity. Therefore, a reduced cognitive capacity
like working memory may lead to the failure of accurate pronoun usage in narrative production.

1.1.2.2 Inhibition.

Another cognitive process that may be related to narrative production is inhibition. Specifically, researchers have speculated that inhibition capacity may be responsible for the maintenance of global coherence in narrative production (Juncos-Rabadán, Pereiro, & Rodríguez, 2005). According to Kemper (2006), based on Hasher, Zacks and May (1999), inhibition serves three functions: (1) to prevent intrusion of irrelevant materials into working memory; (2) to delete irrelevant information from working memory; and (3) to restrain probable responses until their appropriateness is assessed. Therefore, if an individual’s inhibition capacity is inadequate, irrelevant contents will be left unchecked in the process of utterance formulation. As a result, the propositions in the resulting narrative may be repeated or irrelevant, not follow a logical chronological order, or even contradict one another, thus compromising the degree of coherence.

1.1.2.3 Inference.

According to McGinnis, Goss, Tessmer and Zelinski (2008), inference refers to the process in which a reader derives information that is implied by a text and not explicitly stated in order to facilitate comprehension. Generating inferences involves the integration of world knowledge with information that is explicitly stated in the text. Myers and Blake (2008) specifically stated that inferences depend on at least four cognitive operations, which are (1) attending to individual cues of the material from which inferences are being generated; (2) selecting relevant cues; (3) integrating all relevant cues; and (4) associating
cues with previous experiences. They further argued that inference is requisite for the 
generation of the overall theme in narrative production (Myers & Blake, 2008). This is 
especially true when narratives are elicited from visual stimuli such as a single picture or a 
picture sequence. Specifically, inference generation facilitates the accurate conceptualization 
and comprehension of the visual stimuli such that the resulting narrative coherently 
represents the original story in the stimuli.

Evidence that supports the claim that inference generation is essential to the 
comprehension of visual stimuli for narrative production comes from research on the role of 
inference in text processing and comprehension. Researchers have come to the conclusion 
that successful inference generation is crucial to skilled comprehension (McGinnis et al., 
2008). In general, text processing and comprehension is believed to involve mental 
representations at three different levels, namely, surface code, text base, and situation model 
(van Dijk, & Kintsch 1983). The surface code represents the explicit meaning of the text 
without additional derivation of implicit information by the reader. The text base is made up 
of both explicit and inferred information of the text, while the situation model refers to the 
macro-perspective of what the text is about and may contain information such as the setting, 
the characters and the main theme. In order to fully understand a given piece of text, 
complete representations are needed at all three levels. McGinnis et al. (2008) claimed that 
inferences are particularly important in the development of text base and situation model. 
Therefore, if one is unable to generate appropriate inferences, comprehension of the text will 
be compromised. Extending such argument to the processing of the visual stimuli for 
narrative elicitation, which can be argued as the pictorial equivalence of a piece of text, one 
must generate the appropriate inference in order to be able to fully comprehend and
accurately conceptualize the visual stimuli so that the resulting narrative accurately represents the stimuli.

1.1.2.4 Symbol processing.

Symbols are entities that contain no inherent meaning; they are used to represent objects and events in real life. According to the literature in the field of augmentative and alternative communication (AAC), symbols can be classified as transparent, translucent or opaque (Beukelman & Mirenda, 2005, p. 36). Transparent symbols are usually self-explanatory and require very little or no education or training in order to understand their meaning. Some transparent symbols are real objects and photographs of real objects. Translucent symbols are somewhat less self-explanatory than transparent symbols but their meaning can be understood with some education. Examples of translucent symbols include line drawings. Opaque symbols are the least self-explanatory. Their relationship with their referents is usually arbitrary, and thus, meaning of these symbols is the most difficult to grasp among the three types of symbols. Words and Braille are two examples of opaque symbols. Visual stimuli in narrative production are essentially symbolic representation of a given event, using one or more of these classes of symbols. Comprehension and accurate conceptualization of such symbolic representation require functional symbol processing skills, which in turn determine the quality and quantity of the narrative produced. Misconception of these symbols may lead to irrelevance and contradiction in the resulting narrative.
1.2 Narrative in Clinical Assessment

As mentioned in the previous subsection, discourse formulation such as narrative production can be considered as “an intellectual activity that entails not only the understanding and manipulation of linguistic information, but also involves cognitive operations essential to the organization of information” (Ulatowska & Chapman, 1994, p. 29). Therefore, using narrative production in the assessment of cognitive and language disorders in both children and adults can yield a vast variety of information that is invaluable to treatment planning. The following sections review some findings of clinical assessment and studies using narrative production, as well as the methods clinicians and researchers use to analyse the narratives produced by individuals from clinical populations.

1.2.1 Clinical assessment and research using narrative production

According to Ehrlich (1994), the application of discourse study, such as narrative production, to adult clinical populations such as individuals with acquired language disorders properly began with aphasia. Many assessment tools in the field of acquired language disorders have subtests where the patient or client is asked to produce a narrative elicited from some kind of a stimulus. Two famous narrative production subtests are the “Cookie Theft” and the “Picnic” single picture description task from the Boston Diagnostic Aphasia Examination (BDAE; Goodglass & Kaplan, 1972) and the Western Aphasia Battery (WAB; Kertesz, 1982) respectively. Clinicians and researchers have found that individuals with aphasia, especially non-fluent aphasia, tend to produce narratives that are characterized by reduced microstructure (i.e. linguistic forms). This is evident by the simplistic and telegraphic nature of the resulting narratives, which suggests impairment in lexical and
syntactic processing (Ehrlich, 1994; Ulatowska & Chapman, 1994). However, the global features of the narratives are usually relatively preserved (Ehrlich, 1994; Ulatowska & Chapman, 1994).

With the ever increasing recognition and acknowledgement of the importance of extralinguistic and pragmatic variables in communication, narrative production, due to its resemblance to naturalistic speech, has become increasingly popular in the assessment of individuals from clinical populations in which the impairment is primarily cognitive rather than linguistic, as it is in aphasia (Ehrlich, 1994). Two of these clinical populations are individuals with traumatic brain injury (TBI) and right hemisphere damage (RHD). Narratives produced by individuals with TBI tend to have a compromised macrostructure despite the relatively intact linguistic forms. This is possibly due to an impairment of sustained attention and the inability to implement appropriate organizing schemas (Ylvisaker, Szekeres, & Feeney, 2008). Individuals with RHD also tend to experience problems generating macrostructure in narrative production due to their inability to integrate components of a narrative to form a coherent and meaningful whole (Tompkins, Fassbinder, Lehman, & Baumgaertner, 2002). As a result, the resulting narratives tend to be tangential and marked by the absence of the gist and the decreased number of core concepts conveyed (Myers & Blake, 2008).

1.2.2 Analysing narratives

Depending on the goal of the assessment or research, clinicians or researchers can choose to examine various aspects of the micro- and macrostructure of the narratives.
produced by their patients or participants. Several methods and measurements have been
developed in order to investigate these aspects.

1.2.2.1 Lexical measures.

Two major ways clinicians and researchers measure lexical ability are type-token
ratio (TTR) and verb-noun ratio. According to Gordon (2008), TTR is a common measure of
lexical diversity. In its calculation, each occurrence of a word is considered a separate token,
while only the uninflected root of each word is counted as a separate type. For example, the
three tokens *eat, eating* and *ate* would be counted as the same type. TTR is then computed by
dividing the total number of different words (i.e. type) in the language sample by the total
number of words (i.e. token) in the sample. A ratio closer to zero reflects less diversity of
vocabulary, whereas values closer to one reflect greater diversity.

Verb-noun ratio has also been measured in discourse research (e.g. Joanette &
Goulet, 1990; Marini et al., 2005). However, researchers seldom provided justifications for
doing so. One possible reason for examining verb-noun ratio in spontaneous discourse is that
verb use has been found to be differentially more impaired compared to noun use in clinical
studies that employed naming tasks (Mätzig, Druks, Masterson, & Vigliocco, 2009).
Researchers have argued that this may be due to the more complex grammatical status of
verbs (Mätzig et al., 2009). In terms of syntax and semantics, verbs function as predicates
while nouns function as arguments. Each verb consists of a semantic frame in the mental
representation that contains information about the argument structure of the verb, such as the
number of arguments it takes and the semantic role each argument fulfills. Therefore, verbs
have been argued to be inherently more complex linguistically than nouns. Given the
findings from naming tests that verb use tends to be more vulnerable than noun use in the case of acquired language disorders, measuring verb-noun ratio in spontaneous language samples such as elicited narratives allows clinicians and researchers to examine whether a similar trend can be found in more naturalistic speech behaviours.

1.2.2.2 Cohesion and coherence.

According to Joanette and Goulet (1990), cohesion is defined as semantic relations between discourse units (e.g. words, phrases and sentences) that are expressed through specific content or function words. According to Halliday and Hasan (1976), as cited in Brown and Yule (1983) and Duchan (1994), relations among nouns or noun phrases can be expressed by nominal cohesive devices such as reiteration, substitution, ellipsis and referencing. On the other hand, relations among clausal units can be expressed by non-nominal cohesive ties such as additive, adversative, temporal and causal connectives. To examine the cohesion of a given piece of narrative, the appropriate use of these cohesive devices is often measured. Referencing, as a nominal cohesive tie, has received much attention in discourse research. According to Ulatowska et al. (1986), reference is “used to specify the entities to which one’s predications should be linked” (p. 26). Referencing can be achieved by the use of proper nouns, descriptive noun phrases (e.g. “the girl in red”) and pronouns etc. The number of unambiguous references has been used as an indicator of the degree of cohesiveness.

Coherence, on the other hand, refers to the semantic relations that are not expressed through specific use of elements of the discourse (Joanette & Goulet, 1990). Rather, it has to do with the logicality and closeness of how the discourse units relate to each other and the
overall theme of the discourse task at hand. Different coding and scoring schemes have been proposed by clinicians and researchers to examine the degree of coherence of a given piece of discourse. For example, based on Charolles (1978), Joanette and Goulet (1990) proposed and adopted a coding scheme for identifying and analyzing coherence errors in their study that investigated coherence of narratives produced by individuals with RHD. Specifically, they proposed that the following three discourse behaviours represent the degree of coherence of the overall piece of discourse. Non-progression refers to moments in the discourse where an utterance does not lead to the addition of new information or development of the discourse being produced. An example would be the mere repetition of the same utterance, concept or event. Irrelevance refers to moments when an utterance is tangential or clearly off topic. For example, if the topic of a conversation is dinner plans for this evening, a speaker would commit an irrelevance error if he started talking about how much he enjoyed his vacation last summer. Contradiction occurs when one utterance in the discourse clearly contradicts another utterance produced by the same speaker. For example, if the utterance “it is really clear out today” and the utterance “rain is pouring down really hard today” are produced by the same speaker in the same piece of discourse, then it will be marked as a contradiction error.

1.2.2.3 Content richness.

To investigate the content richness of narratives, a proposition analysis is often needed. Kintsch (1974) and Turner and Greene (1977) have done extensive research in this area. According to these researchers, a proposition can be defined as an idea unit, which consists of word concepts. One of these word concepts serves as the relation, either as a predicate, modifier, or connective. The other ones serve as arguments of the proposition. All
propositions of a given piece of discourse are then connected to form a text base. The goal of proposition analysis, then, is to identify and extract all the propositions from each utterance in order to create a list of connected idea units that represent the core concepts of a given piece of discourse.

Once all propositions of a piece of discourse have been identified, researchers then adopt different ways to measure content richness. One of the more widely used analyses is to calculate the number of propositions or content units relative to the total number of words in a piece of discourse. Kemper and Sumner (2001), based on Kintsch and Keenan (1973), and Kemper, Thompson and Marquis (2001) called this measure propositional density (P-density). According to Kemper and Sumner, P-density is a measure that examines “the consistency with which ideas are expressed succinctly or not” (p. 319). Another popular way to examine content richness that is specific to narrative studies is to calculate the number of main or expected propositions each participant produces in his or her narrative for a particular story. The main or expected propositions for each story are usually predetermined in a pilot study (Marini et al., 2005), or by the investigators (Joanette & Goulet, 1990; Pratt et al., 1989; Wright, et al. 2005). In the former condition, researchers usually develop the set of main propositions by including these propositions produced by the majority (e.g., 75% to 80%) of their pilot study participants.

1.3 Age-related Differences in Narrative Production

Psycholinguistic studies and studies that investigated age-related difference in language use have shown a change in discourse behaviours at various domains of language as
age increases. The following sections examine findings of the current literature regarding the effects of age on aspects of language use that may impact narrative production.

1.3.1 Age and lexical processing

1.3.1.1 Lexical diversity and access.

Various researchers have investigated the effects of age on lexical diversity as measured by type-token ratio (TTR). Nef and Hupet (1992) conducted a study that looked at the relationship between age and spoken and written discourse production. Healthy adults between the ages of 40 and 90 were recruited to participate in the study, where they were asked to produce language samples in picture description and narrative production tasks. Results showed that unlike other linguistic measures such as the use of complex sentences, TTR remained the same across age groups. Kemper and Sumner (2001) also measured TTR of discourse produced by young and older adults. Specifically, they recruited 100 young adults between the ages of 18 and 28, and 100 older adults between the ages of 63 to 88 to participate in their study. All participants were asked to produce a five-minute description of an influential person or event in their life. Like Nef and Hupet, they did not find a difference in TTR between the two age groups. Instead, TTR was found to be higher in the older group of participants.

While studies on lexical diversity did not find an age-related disadvantage for the older adults, studies on lexical access, on the other hand, yielded mixed evidence regarding older adults' ability to retrieve lexical items compared to that of younger adults. Seeking to investigate the relationship between aging and lexical access of nouns, Evrard (2001) ran a study that involved administering a picture-naming test on three groups of participants,
namely, young (18 to 34), middle-aged (35 to 54) and older (55 to 75) participants. It was found that although the older participants experienced more difficulty retrieving proper nouns, no age difference was found in terms of the lexical access of common nouns. The authors attributed the greater difficulty with accessing proper nouns experienced by the older participants to the lack of inherent meaning for proper nouns. Because retrieval of a word was argued to depend on the access to the semantic properties of the referent, the lack of semantic properties for proper nouns made its retrieval more challenging. Cerella and Fozard (1984) similarly compared lexical access between younger and older adults. The mean age for the young and the older group was 24.1 and 73.1 respectively. The researchers found that lexical access was unaffected by age.

Other studies have found evidence contrary to these findings. Schaie and Willis (1993) examined the effects of age on cognition and language use, using a series of psychometric tests. 1628 adults were recruited to participate in this study. Participants were divided into 9 different age groups, with mean age ranging from 29 years to 88 years. In terms of language, tests such as vocabulary recognition and the Educational Testing Service (ETS) vocabulary tests were used to measure verbal ability, while a verbal fluency test and a memory test for words were administered to measure verbal memory. Verbal ability was found to remain stable throughout middle-age until the age of 67, where the start of a gradual decline was noted. As for verbal memory, a gradual decline starting at 39 years old all the way to the elderly ages was noted. Juncos-Rabadán and Iglesias (1994) also found cross-linguistic evidence that supports a decline in lexical ability with age. They administered the Bilingual Aphasia Test on 840 adults of 14 different ethnicities between the ages of 50 and 91. Performance on the synonym and antonym naming subtests were found to decline with
increasing age, which the researchers concluded to be due to a decline in lexical access. If lexical access does decline with age, then lexical diversity in discourse, which relies on successful and prompt lexical access, might be expected to decline as age increases.

The observation from the above studies that lexical diversity is not negatively impacted by age while increasing age may adversely influence lexical access is intriguing. This is because unimpaired lexical access is arguably a critical factor that contributes positively to lexical diversity in a given piece of discourse. The more smoothly and efficiently one is able to access lexical items, the more likely that one would be able to use a greater variety of vocabulary, hence increasing lexical diversity. If lexical access does not remain stable as age increases, it is then curious that lexical diversity does not also decline with age.

The study by Kemper and Sumner (2001) described earlier offers some insight into explaining these apparently contradicting findings. They found that lexical access measured during formal testing that elicit single-word responses may be somewhat less related to lexical diversity measured during more naturalistic speech than one might assume. Specifically, in addition to measuring TTR in narrative production, several vocabulary tests, including the Shipley Vocabulary Test, the Wechsler Adult Intelligence Scale – Revised (WAIS-R) Vocabulary Test, and the Peabody Picture Vocabulary Test (PPVT) were administered. The authors then calculated correlations between the TTR of the language samples and the score for each of these psychometric tests. Only weak-to-moderate positive correlations were found between TTR and the results of psychometric testing for both the young adult and older adult group. This suggested that even though performance on
psychometric tests that examine lexical access may decline with age, the same trend may not be evident in lexical diversity measures derived from language samples.

1.3.1.2 Verb-noun ratio.

Like lexical access, results of research on the relationship between age and verb-noun ratio appear to be inconclusive. In Marini et al.'s (2005) study, which looked at the effects of age and the use of single picture vs. picture sequence, verb-noun ratio of the narratives produced by participants in both stimulus conditions remained consistent across all age groups (20 to 84). This finding does not appear to agree with findings from studies that revealed dissociation between performance on noun-naming and verb-naming tests. For example, Barresi, Nicholas, Connor, Obler, and Albert (2000) conducted a study that investigated the effect of age on the ability to name nouns vs. verbs. The Boston Naming Test (BNT) was used for the noun-naming task, while the Action Naming test (ANT) was used for the verb-naming task. Both tests are essentially a confrontation naming task, where participants are shown a picture and asked to name the noun (in the BNT) or verb (in the ANT) for each test item. Thirty-nine adults between the ages of 50 and 79 participated in the study. Results of the study showed that as age increased, failures to name nouns on the BNT increased. However, the same pattern was not observed in performance on the ANT. In other words, it appeared that aging affected noun but not verb naming. If this were the case, one would anticipate an increase in verb-noun ratio in narratives produced by older as opposed to younger adults, which was not observed in Marini et al.'s (2005) study.

It is important to note, however, that the discrepancy between the findings of the above studies must be interpreted cautiously. Firstly, recall that in the review of literature on
age effects and lexical processing, it was noted that results of naming or psychometric testing may not always correlate with language measures obtained from spontaneous speech. The same can be argued in the case of verb and noun retrieval in a naming task vs. discourse production. Specifically, verb and noun retrieval in a naming task are not exactly the same as the use of verbs and nouns in discourse tasks like narrative production. For example, in the former task, each lexical item is activated independently and purely by visual input, whereas in the latter task, elements of the discourse environment such as context and topic, as well as the linguistic environment of the continuous speech may have cuing effects on verb and noun retrieval to a point where ceiling effects occur. As a result, even though noun retrieval appeared to be more negatively affected by aging than verb retrieval in the context of confrontation naming, such a pattern may not be observed in narrative production. Secondly, Mackay, Connor, Albert and Obler (2002) pointed out that item difficulty of the BNT and the ANT may not be the same in the first place. Any difference found between the performance on the BNT and the ANT, therefore, could simply be an artifact of such differences between the two tests. To further explore the matter, Mackay et al. (2002) also conducted a study that investigated the effect of aging on noun and verb retrieval, using the BNT and the ANT. However, instead of comparing their participants’ performance on the whole BNT and the whole ANT, they underwent a matching procedure in which they selected a subset of items from each test that were judged to be statistically equal in terms of item difficulty. Participants’ performance on these two subsets was compared. Results showed that when item difficulty of the two tests was counterbalanced, no divergence in performance on noun vs. verb naming with increasing age was observed.
Interestingly, Earles and Kersten (2000) found yet another pattern of discrepancy between noun and verb retrieval. A younger group (18-32) and an older group (60-82) of participants took part in this study, where they were presented with verb-noun pairs and were then asked to recall them without cues. It was found that verb recall was significantly more difficult than noun recall for older individuals. If one closely compares Earles and Kersten’s and Marini et al.’s (2005) study, it becomes clear that their results do not really contradict each other. In Earles and Kersten’s study, the task that they chose was an episodic memory task, where participants were asked to recall words previously presented. In Marini et al.’s study, however, where noun and verb use in narratives was examined, phonological retrieval of lexical items from semantic memory was tested. As a result, the greater difficulty in retrieving verbs in Earles and Kersten’s study can be attributed to a decline in episodic rather than semantic memory for verbs in elderly. This is in line with the findings in various studies on cognitive decline such as Alzheimer’s disease and amnestic mild cognitive impairment that semantic memory appeared to be more robust and resistant against deterioration than episodic memory (Hopper and Bayles, 2008; Murphy, Troyer, Levine & Moscovitch, 2008). When verb retrieval involves semantic memory processes, as in naming or use in discourse, the effects of age on successful retrieval compared to nouns should be minimal, as exemplified by Marini et al.’s study.

1.3.2 Age and cohesion and coherence

Another area that is commonly examined by studies that look at age differences in discourse production is cohesion and coherence. The following sections examine the effects of age on cohesion as measured by reference use and on coherence. The various possible mechanisms through which age may affect cohesion and coherence are also explored.
1.3.2.1 Cohesion.

As discussed in section 1.2.2.2, referencing is a major cohesive device in discourse production such as narrative. Various linguistic studies have explored the effects of age on the quality of referencing in discourse production. Ulatowska et al. (1986) invited 51 nuns between the ages of 27 and 92 to participate in their study. The 51 nuns were identified because of their homogenous educational and socio-economic background. They were asked to complete several cognitive tasks and discourse production tasks, including simple and complex story retell, conversation, and self-generated and procedural narrative production. Results of the study revealed that difficulties with referencing began to emerge in the young elderly years and became very prominent by the age of 76. Similar results had also been found by Marini et al. (2005) in their aforementioned study and by Pratt et al. (1989) and Juncos-Rabadán et al. (2005) in their narrative studies on adults between the ages of 18 and 87, and 41 and 91 respectively. Specifically, in all three studies, increasing age was found to negatively affect referencing, leading to a significant number of missing or ambiguous referents in narratives.

Researchers have attempted to uncover the underlying factors that contribute to the decline of narrative cohesion as measured by clear and accurate referencing associated with age. In Ulatowska et al.’s study (1986), they administered several cognitive tests in addition to different discourse production tasks. Similar to referencing, performance on all cognitive tasks was significantly negatively correlated with age. Therefore, the researchers speculated that the decline in the use of referencing as a cohesive device may be attributable to a general decline of cognitive ability as age increases.
Among the various cognitive abilities, working memory has received much attention in research examining age-related declines of discourse cohesion. Recall that in section 1.1.2.1, working memory has been shown to be closely related to the quality of referencing in narrative production (Pratt et al., 1989). Various research studies have shown that as age increases, working memory capacity reduces. Salthouse, Kausler, and Saults (1988) conducted a study that investigated age differences in performance on memory tasks that separately examined the visual-spatial and the phonological loop slave system of working memory. 362 adults between the ages of 20 and 79 participated in the study. The matrix memory task was administered where participants were presented with a five-by-five matrix with 25 letters for three seconds. Seven letters were marked as target items. In the visual-spatial condition, participants were asked to recall the location of the target items regardless of what the letters actually were. In the verbal condition, participants were asked to recall the target letters in any order. An age-related decline of performance in both the visual-spatial and the verbal condition was noted, suggesting that as age increases, the capacity of the two slave systems in working memory becomes increasingly compromised.

Borella, Carretti, and De Beni (2008) looked at the effect of advancing age on visual and verbal working memory (i.e., the central executive plus one of the slave systems). They recruited 304 adults between the ages of 20 and 86 and asked them to complete one visual working memory task and two verbal working memory tasks. They found that age was negatively correlated with the performance on these working memory tasks. In light of the results in Salthouse et al.'s (1988) study, such a finding was not surprising since the two slave systems were an integral part of visual-spatial and verbal working memory respectively. Decline in these slave systems would inevitably lead to a decline in working
memory. Findings of this study were also consistent with the results of Park, Lautenschlager, Hedden, Davidson, Smith and Smith's (2002) study, which involved the completion of visual and verbal working memory tasks by 345 adults between the ages of 20 to 92. In this study, it was found that working memory processes declines as age gradually increases regardless of task modality (i.e., visual-spatial vs. verbal).

Therefore, if working memory declines with age, and if poor working memory is correlated with referencing errors, one may speculate that the age-related deterioration of reference use may be at least partially due to a decline in working memory.

1.3.2.2 Coherence.

While the current literature reviewed unanimously suggested a reduction in discourse cohesion with aging, studies have come to different conclusions regarding whether there is also a negative age effect on discourse coherence. Mackenzie (2000) conducted a study in which the effect of age on coherence, as measured in terms of the degree of verbosity, topic maintenance, and the amount of relevant content, was examined in the context of conversation and picture description. Results showed that the older elderly group (75 to 88) was more inclined to verbosity and off-topic speech in conversation. Older elderly were also found to be less efficient in conveying information in the picture description condition. Putting this into Joanette and Goulet’s (1990) coding scheme (see section 1.2.2.2 above), these findings suggest that older individuals may be more prone to committing non-progression and irrelevance errors, as evident by an increase in verbosity and difficulty maintaining topic respectively with age. Juncos-Rabadán et al. (2005) found comparable results in their study where participants were instructed to produce narratives based on single
pictures. Coherence was measured in terms of the faithfulness of content (i.e. whether the content expressed corresponded to that of the story as represented in the pictures) and the amount of irrelevant content. They found that the faithfulness of content reduced over age while the amount of irrelevant content increased. In terms of Joanette and Goulet’s (1990) scheme, this suggests an increase in irrelevance error with advancing age. They attempted to explain these results with Gold and Arbuckle’s (1995) and Pushkar et al.’s (2000) view on verbosity and aging. Specifically, they speculated that the reduction in content faithfulness could be explained by an age-related decline in cognitive inhibitory function, which, as discussed in section 1.1.2.2, is essential to maintaining coherence during narrative production.

Other researchers, however, have not found any effects of age on coherence. Beaudreau, Storandt and Strube (2006) examined the coherence of narratives produced by 24 young (18 to 21) and 24 older (56 to 91) adults. Three narratives were elicited respectively by three pictures that conveyed different emotions. The amount of off-topic speech was recorded as a measure of coherence. The researchers did not find an effect of age on content relevance in all narratives produced by participants. In Marini et al.’s (2005) narrative study, coherence was examined in terms of irrelevant and misinterpreted content with respect to the overall gist of the story. Like Beaudreau et al., they did not find a significant effect of age on the number of coherence errors.

If one compares the above studies carefully, one may realize that the discrepancy among the findings in these studies could be partially due to the different discourse tasks involved. In particular, Mackenzie (2000) and Juncos-Rabadán et al. (2005) elicited narratives from their participants by using a single picture in their respective studies. On the
other hand, Marini et al. (2005) employed a picture sequence description task to elicit narratives. The former task can be argued as more cognitively challenging since the participants were asked to formulate their own story based on just one picture. In the latter task, the formulation of the story line was constrained by the picture sequence. Such constraint might actually help guide the participants to produce contents that were on-topic and relevant to the story. In other words, the picture sequence formed an extrinsic inhibitory force that helped to keep irrelevant content in check. As a result, the narrative task became less cognitively demanding (e.g. less intrinsic inhibition required) such that both young and older adults could produce narratives with few coherence errors. Therefore, no age difference in terms of coherence was observed between the younger vs. older adults in studies that used picture sequence to elicit narratives.

The discrepancy among the findings in these studies may also be attributed to the difference in the age of the older groups of participants. For example, in Mackenzie’s (2000) study where an age difference was found in coherence, the age range of the oldest elderly group was 75 to 88. However, Beaudreau et al.’s (2006) study where no age difference was found in coherence, the average age of the older adults was 72.13. Therefore, one may argue that no age difference in coherence was observed in the latter study because the older participants in that study had not reached the developmental stage where decline in coherence in narrative production emerges.

1.3.3 Age and content richness

The third popular area of discourse analysis regarding age difference is content richness and whether informativeness in discourse production decreases with age. Research
studies have come to very similar conclusions that age has a negative impact on content richness of the discourse produced. Kemper et al. (2001) found in their longitudinal narrative study that P-density started to decline between the ages of 74 and 78. This result was very compatible with Kemper and Sumner's (2001) findings that a lower P-density was observed in the narratives produced by their older participants. Studies that measure the number of expected main propositions produced for a given narrative to explore content richness also arrived at very similar results. Duong and Ska (2001) conducted a study that investigated narrative production by normal healthy adults using activities such as single-picture and picture-sequence description. Adults in the older groups (65 to 84 years) produced significantly fewer expected main propositions than those in the younger groups, although the authors did not specify the age range of the younger groups. Wright et al. (2005) also compared content richness of narratives produced by younger (21 to 28 years) and older adults (57 to 83 years) using similar activities, namely, single-picture and picture-sequence description, and came to the same results. In the aforementioned studies by Marini et al. (2005) and Pratt et al. (1989), where both groups of researchers looked at narrative production by younger and older adults, it was found in both studies that the number of expected main propositions produced by participants decreased with advancing age.

Various hypotheses to account for the decline of content richness with age in narrative production have been postulated. For example, just like the age-related decline of coherence discussed above, Kemper (2006) suggested that an age-related deficit in inhibition could be responsible for the decline of language abilities, including P-density. Such postulation is reasonable, considering how P-density is calculated. When inhibition becomes compromised over age, irrelevant and redundant materials are more likely to be included in
utterances produced. However, in the calculation of P-density, each proposition is only counted once. That is, any repeating propositions will not be included in the calculation, yet the repeated materials will be added to the total number of words produced. Mathematically speaking, when repetitions occur in a piece of discourse, the numerator (total number of propositions) of the P-density equation remains constant, while the denominator (total number of words) increases. This will yield a smaller P-density value. Therefore, with reduced inhibition as age increases, content richness, as measured by P-density, decreases.

Other researchers have also found a correlation between lexical ability and content richness of narratives produced by healthy adults. Cheung and Kemper (1990), as cited in Kemper and Sumner (2001), ran a study that explored the relationships among linguistic measures gathered from language samples produced by younger and older adults. They found that verbal ability measured solely by the WAIS-R Vocabulary test predicted P-density of narratives. Recall in the discussion on lexical access in section 1.3.1.1 that although lexical diversity appeared to remain stable across ages, several studies have found evidence for an age-related decline in performance on psychometric vocabulary tests. Kemper and Sumner further explored the relationship between vocabulary and proposition density by including verbal fluency and other processing tasks in their narrative study. However, they found that proposition density appeared to be more of an indicator of processing efficiency.

Similarly some other researchers proposed possible underlying reasons for the reduction in content richness with age. For example, Duong and Ska (2005) attributed the decrease in the number of expected main propositions produced by their older participants to a decline in the ability to organize discourse and the ability to conceptualize the visual stimuli. Such age-related deficits prevented the older participants from producing narratives
that are rich in content. In studies where story retell were used as the narrative production
task, such as in Pratt et al.'s (1989) study, researchers have attributed the decline in content
richness in narrative produced by older participants to a decline in working memory capacity
with age. Pratt et al. specifically narrowed the focus down to the process of encoding in
working memory, and stated that older participants may be experiencing greater difficulty in
encoding story materials, thus depriving themselves of the opportunity to produce an
informative narrative in a story retell task.

1.4 The Use of Different Visual Stimuli in Discourse Elicitation

In section 1.1.1.1, it was shown that narratives can be elicited by using various
procedures. Researchers have conducted studies which are highly relevant to clinical practice
on the effect of using different kinds of visual stimuli and/or tasks on narrative production.
Specifically, in order to be able to compare the performance in narrative production across
populations, especially clinical populations, the task that is used to elicit the narratives
becomes of critical interest. The various media that are frequently used in narrative
production tasks and examined by researchers include single pictures, picture sequences and,
less frequently, video. In a single picture task, participants are usually asked to generate a
story based on a single picture. In a picture sequence or a video task, participants are asked to
retell the story presented to them through the picture or video stimulus. This section explores
some potential differences among the use of the different visual materials in narrative
production. The current literature on how the use of different visual stimuli, namely, single
picture vs. picture sequence and picture sequence vs. video may influence narrative
processing is also reviewed.
1.4.1 Potential differences among single picture, picture sequence and video

Firstly, the demand on working memory capacity imposed by a single picture vs. a picture sequence stimulus is potentially different. In the former condition, as described earlier, participants are asked to generate an entire story based on just one single picture or scene. In the latter condition, the entire story is usually presented to the participants through the picture sequence and all they are asked to do is to retell the story. Given this potential difference, a single picture task arguably imposes a greater demand on the cognitive system, including working memory, than the picture sequence task does. The demand on working memory imposed by a picture sequence vs. a video stimulus is also arguably different. In a picture sequence condition, participants can be allowed to look at the visual stimulus while they are producing the narrative. However, they cannot be allowed to do so in the video condition. This is because simultaneously viewing the visual stimulus while producing the narrative would implicitly create a time constraint on the participants’ narrative production. On the other hand, pausing the video to eliminate this time constraint also would not suffice because that would change the video to a condition that resembled a picture sequence. Therefore, participants should not be allowed to look at the visual stimulus in the video condition while they are producing the narrative. This, in turn, could place a greater demand on the participants’ working memory capacity in the video condition, however, as they would need to store the entire story line in memory while simultaneously formulate utterances to verbally express it.

Secondly, recall that in section 1.1.2.3, it was discussed that inference generation is a crucial mental process in narrative production tasks. The demand on inference generation is arguably higher in a single picture task than in a picture sequence task. In the former
condition, participants must make multiple inferences about the event(s) happening in the story based on just one single scene before they can construct a narrative. In the latter condition, more information about the event(s) of the story is given to the participants through the sequence of pictures in the stimulus, thus requiring them to make fewer inferences about the story. Such a difference in inference generation between single picture and picture sequence may also be argued to exist between picture sequence and video. In the video condition, there is a continuous presentation of information in an uninterrupted manner. When the stimulus is represented by a picture sequence, however, interruptions of the presentation of information occur between the picture frames of the sequence. Because no information is presented at these gaps, participants need to bridge these gaps by utilizing inferences in order to make sense of the whole picture sequence. Therefore, it is speculated that the picture sequence condition imposes a greater demand on inference generation than the video condition.

Thirdly, the amount of symbol processing required by a single picture or a picture sequence vs. a video is different. As discussed in section 1.1.2.4, because visual stimuli used in narrative elicitation are essentially symbolic representations of one or multiple events, symbol processing was shown to be critical to comprehension and accurate conceptualization of the story. Given the classification scheme of symbols discussed in section 1.1.2.4, video can be argued as transparent while picture (single picture or picture sequence) is more of a translucent symbolic representation. In other words, although the meaning of a picture stimulus should be fairly easily accessed by the picture reader, it is still more opaque than that of a video stimulus, whose meaning is even more accessible and thus, requires less amount of symbol processing.
1.4.2 Single picture vs. picture sequence

The following sections review some of the findings of how single picture vs. picture sequence influence cohesion, coherence and content richness in narrative production.

1.4.2.1 Cohesion.

Liles et al. (1989) examined the use of picture sequence vs. single picture on narrative production by normal healthy adults and adults with closed head injury (CHI). In the picture sequence condition, participants were shown a story represented by 19 picture frames, and were asked to retell the story afterwards. In the single picture condition, participants were asked to generate a story based on a single picture. The researchers found that in terms of cohesion, both the normal healthy group and the CHI group produced more cohesive ties that are clear and complete in the picture sequence condition. They attributed the differences between the two tasks to the possibility that the single picture condition was more cognitively demanding than the picture sequence condition. In the aforementioned study by Duong and Ska (2001), where participants were asked to produce one narrative for a single picture and a picture sequence stimulus respectively, it was found that the picture sequence condition appeared to facilitate the use of a greater number of transitional markers for both younger and older adults.

1.4.2.2 Coherence.

Similar to cohesion, the use of different visual stimuli has also been shown by research to influence the degree of coherence in narrative production. In Marini et al.'s (2005) study, each participant was asked to produce three narratives: one elicited from a single picture, and two from two different picture sequences. Global coherence errors, which
are operationally defined as instances where an utterance was unrelated or conceptually incorrect with respect to the overall theme of the story, were coded for each narrative. Similar to the findings on cohesion from Liles et al.’s (1989) and Duong and Ska’s (2001) studies, the degree of global coherence was significantly lower in narratives elicited from the single picture condition for all age groups.

The result of Marini et al.’s study can be explained by the potential difference between single picture and picture sequence in terms of inference generation postulated above. That is, a single picture task arguably requires more inference generation than a picture sequence task. As discussed in section 1.1.2.3, inference is crucial to successful conceptualization and comprehension of the visual stimuli. Therefore, if a greater demand is placed on inference processes, as in the single picture condition, it is more likely for the participants to misconceptualize or misunderstand the story represented by the visual stimulus. As a result, the resulting narrative may consist of more irrelevant, repetitive or contradicting materials, thus decreasing the degree of coherence of the narrative.

1.4.2.3 Content richness.

In Liles et al.’s (1989) study mentioned above, the researchers also compared the number of complete episodes of events elicited in the single picture vs. picture sequence conditions. They found that all their participants (with or without a closed head injury) produced more complete episodes of events in the picture sequence condition compared to the single picture condition.

In the aforementioned study conducted by Wright et al. (2005), where the effects of age and using single picture vs. picture sequence on content richness in narrative production
were examined, it was found that normal healthy adults tended to perform better in terms of content richness in the picture sequence condition as opposed to the single picture condition. Specifically, regardless of the participants’ age, they tended to produce more relevant main propositions in the picture sequence condition. Capilouto, Wright and Wagovich (2006) conducted a follow-up study to Wright et al.’s (2005) study, comparing content richness of narratives produced by neurologically intact individuals vs. individuals with aphasia. Eight participants with mild to moderate aphasia and eight normal healthy participants were recruited for the study. They were each asked to produce narratives from two single picture and two picture sequence stimuli. The total number of main events in the narratives elicited from each stimulus was measured. Results of the study closely resembled those from the previous study, namely, participants of both groups produced significantly more main propositions in the picture sequence conditions.

Similar to the reduced degree of coherence observed in narratives elicited from a single picture discussed in the previous section, the findings of these studies may also be attributed to the greater demand imposed on inference generation by the single picture condition. Because of the smaller demand on inference generation in the picture sequence condition, misconception of the visual stimulus is less likely to occur. As a result, more main events and propositions were produced in narratives elicited from a picture sequence.

1.4.3 Picture sequence vs. video

Research on the effects of using picture vs. video as elicitation materials in narrative production was found to be extremely limited. A literature search using different combinations of the key words “picture”, “video”, “visual stimuli, “narrative”, and “story-
telling” was carried out in three databases of published research, namely, PsycINFO, Linguistics and Language Behavior Abstracts (LLBA), and Cumulated Index to Nursing and Allied Health Literature (CINAHL). No published study was found on the effects of the use of pictures vs. video as visual stimuli on narrative production by adults. Having said that, studies that examined narrative production by children using pictures vs. video as stimuli, as well as studies that examined the use of videos and pictures in second language acquisition may provide insights into how these two stimulus modalities may impact narrative production in adults.

1.4.3.1 Visual stimuli and narrative production in children.

As in the case of adults, the effects of using different visual stimuli on narrative production in terms of content richness have also been studied with children. Verhallen, Bus and de Jong (2006) conducted a study that investigated narrative comprehension and production of kindergarten children who were at risk for school failure, using static pictures vs. multimedia video. This study took place in the Netherlands. Sixty kindergarten children, who were identified as at risk for school failure based on their socio-economic status and the fact that Dutch was their second language, participated in the study. All children were exposed to a story in Dutch either in the form of static picture sequence plus oral narration or multimedia video plus oral narration. The children were then asked to retell the story as completely as they could with the aid of the visual stimuli depending on which condition they were in, with the oral narration turned off in both conditions. The number of main action events and implied events were coded for each narrative produced. Results of the study showed that children produced more complete retellings in the video condition. Specifically, children in the video condition produced significantly more implied events that referred to
causal and enabling relationships between groups of events. The researchers attributed such findings to the speculation that video was more effective in assisting the children in making inferences. This was consistent with the speculation in section 1.4.1 that picture stimuli put a greater demand on inference generation than video stimuli.

1.4.3.2 Pictures and videos in second language acquisition.

Research on using different visual stimuli in assisting children in acquiring a second language acquisition may provide insight into how visual stimuli can affect content richness in narrative production. Hanley, Herron and Cole (1995) explored the use of video and narrative vs. picture and narrative as an advance organizer for learning in foreign language elementary school classrooms. An advance organizer can be defined as “the process of linking the unfamiliar to what is already known by the learner” (p. 57). In other words, this study was conducted in order to explore how pictures and videos act as an assistive tool to facilitate comprehension and retention of a piece of text in a foreign language. Sixty-two fifth-graders who were learning French at school participated in this study. Students were exposed to either a picture and an oral narrative or a video and a narrative before each French class as an advance organizer in aiding the learning process. During class, the written materials, on which the participants were tested on immediately after each class and at the end of the experiment, were taught. The researchers found that students in the video and narrative condition performed significantly better in immediate tests after each class and in the final test at the end of the study. They attributed this finding to the tighter “contextual fit” between the video advance organizer (as opposed to the picture advance organizer) and the written passage. That is, the video was more successful in helping the students to
contextualize the subsequence written materials. They therefore concluded that video was more effective in aiding comprehension in second language acquisition.

1.5 Summary of Literature Review

In section 1.3 and section 1.4, literature on the effects of age and task on narrative production was reviewed. In summary, age did not appear to have any effect on lexical diversity measured by type-token ratio (TTR) and on verb-noun ratio. However, age appeared to have a negative effect on the degree of cohesion and content richness. The decline in cohesion may be associated with an age-related decline in working memory capacity, while the decline in content richness may potentially be due to age-related changes in inhibition and the ability to conceptualize from visual stimuli that elicit the narratives. With regard to the effect of age on coherence, findings from research studies were inconclusive. Such discrepancies in findings could be attributed to the different elicitation materials and tasks used and the different age range of the older participants in these studies.

Regarding the effects of different visual stimuli on narrative production, studies found that picture sequence tended to elicit narratives that were more cohesive, coherent, and richer in content regardless of age. Research studies that looked at narrative production by children and the use of picture vs. video in second language acquisition were also reviewed. Results of these research studies suggested that video was more successful than picture sequence in facilitating better performance in story-retelling. Video was also shown to be a more effective advance organizer than pictures in assisting children to acquire a second language.

Based on the previous review, a logical next step for research is to further explore the effects of age on narrative production, using different kinds of visual stimuli. The use of
picture sequence vs. video in narrative elicitation is of particular interest because of the relatively scarce research in this area, particularly with the adult population, and the possibility that different types of visual stimuli could have different effects on narrative production across age groups.

1.6 Research Hypotheses

This study is designed to explore the effects of age and the use of different visual stimuli, namely, video vs. picture sequence, on narrative production by younger and older normal healthy adults in terms of lexical processing, cohesion and coherence, and content richness. Based on the literature reviewed in Chapter 1, the following hypotheses are postulated:

1. Age will not have an effect on:

   a. lexical diversity
   
   b. verb-noun ratio
   
   c. coherence

As discussed in section 1.3.1.1, lexical diversity, as measured by TTR is unlikely to be affected by age in narrative production. The discussion in section 1.3.1.2 suggested that verb-noun ratio is also unlikely to be affected by age in narrative production. While mixed evidence was found regarding whether age influences coherence on narrative production, the discussion in section 1.3.2.2 suggested that studies that found an age effect on coherence might have employed a more cognitively demanding elicitation task (single picture description) than studies that did not find such an effect (picture sequence description), leading to the inconclusive findings. Because the current study uses picture sequence and
video, which were argued to be less cognitively demanding for narrative elicitation, no age difference in terms of coherence is anticipated in this study.

2. Age will have a negative effect on:

   a. cohesion
   b. content richness

   In section 1.3.2.1, it was demonstrated that as age increases, the ability to use clear references in connected speech such as narrative declines, which may be at least partially due to a decline in working memory with age. Section 1.3.3 showed that aging also appears to negatively influence content richness in narrative production.

3. Participants, especially older adults, will commit more cohesion errors in the video condition than in the picture sequence condition.

   Recall that in section 1.1.2.1 and section 1.3.2.1, it was shown that working memory is essential to the use of clear references as a cohesive device in narrative production. Given the speculation in section 1.4.1 that a video stimulus imposes a greater demand on working memory capacity than a picture sequence stimulus in narrative production, it is predicted that participants will make more referencing errors in the video condition. This is anticipated to be especially evident for the older participants due to the age-related decline in working memory discussed in section 1.3.2.1.

4. In the picture sequence condition, participants will:

   a. commit more coherence errors
   b. produce narratives that are less rich in content
Hanley et al.'s (1995) study showed that video was more effective in facilitating children's comprehension in a second language. Verhallen et al.'s (2006) study demonstrated that using video, as opposed to picture sequence, in a story retell task led to the production of more complete narratives. Also given the speculation in section 1.4.1 that a picture sequence stimulus imposes a greater demand on inference generation and symbol processing, it is anticipated that individuals in the video condition will produce narratives that are more coherent and richer in content than in the picture sequence condition in the current study.
2. Methods

This chapter details the methods used to explore the research questions. Specifically, the demographic information of the participants, the tools used to further characterize the participants, and the method of recruitment are discussed. This is followed by a detailed explanation of the choice and design of the narrative production experimental tasks. The procedures by which participants completed these tasks, and the analyses of the collected data, are also described.

2.1 Participants

2.1.1 Demographic information

Two groups of participants, a younger group and an older group, were recruited for this study. Participants in the younger group (n = 20) had a mean age of 27.3 (SD = 3.93, range = 20 to 36). They were recruited through posters, email notices and by word of mouth circulated within student organizations at the University of British Columbia and the broader community. Participants in the older group (n = 20) had a mean age of 69.6 (SD = 5.02, range = 63 to 81). They were recruited through posters circulated among seniors' organizations in the community and word of mouth. All participants used English as their primary language in their daily living and did not have a previous or current diagnosis of cognitive or language impairment. Because previous research did not demonstrate a gender difference in narrative production (Mackenzie, 2000), gender of participants in both groups was not controlled for in this study. The demographic information about the participants is shown in Table 2.1. The use of human participants for the current study was approved by the Behavioural Research Ethics Board of the University of British Columbia. All participants
provided written informed consent prior to commencing their participation in the study. The ethics certificate and the informed consent form can be found in Appendix A and Appendix B respectively.

Table 2.1 Demographic information about participants

<table>
<thead>
<tr>
<th>Subject</th>
<th>Young</th>
<th>Old</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age</td>
<td>Age</td>
</tr>
<tr>
<td></td>
<td>Education</td>
<td>Education</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>Gender</td>
</tr>
<tr>
<td>1</td>
<td>36</td>
<td>81</td>
</tr>
<tr>
<td>2</td>
<td>27</td>
<td>71</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>67</td>
</tr>
<tr>
<td>4</td>
<td>27</td>
<td>68</td>
</tr>
<tr>
<td>5</td>
<td>33</td>
<td>65</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>70</td>
</tr>
<tr>
<td>7</td>
<td>25</td>
<td>77</td>
</tr>
<tr>
<td>8</td>
<td>25</td>
<td>66</td>
</tr>
<tr>
<td>9</td>
<td>28</td>
<td>71</td>
</tr>
<tr>
<td>10</td>
<td>24</td>
<td>67</td>
</tr>
<tr>
<td>11</td>
<td>29</td>
<td>65</td>
</tr>
<tr>
<td>12</td>
<td>26</td>
<td>75</td>
</tr>
<tr>
<td>13</td>
<td>31</td>
<td>65</td>
</tr>
<tr>
<td>14</td>
<td>23</td>
<td>67</td>
</tr>
<tr>
<td>15</td>
<td>31</td>
<td>70</td>
</tr>
<tr>
<td>16</td>
<td>25</td>
<td>80</td>
</tr>
<tr>
<td>17</td>
<td>23</td>
<td>68</td>
</tr>
<tr>
<td>18</td>
<td>26</td>
<td>67</td>
</tr>
<tr>
<td>19</td>
<td>32</td>
<td>69</td>
</tr>
<tr>
<td>20</td>
<td>30</td>
<td>63</td>
</tr>
<tr>
<td>Mean</td>
<td>27.3</td>
<td>69.6</td>
</tr>
<tr>
<td>SD</td>
<td>3.93</td>
<td>5.02</td>
</tr>
<tr>
<td>Range</td>
<td>20-36</td>
<td>63-81</td>
</tr>
</tbody>
</table>

2.1.2 Working memory span

As discussed in the previous chapter, working memory was potentially a co-variant of any differences found in narrative production between the two groups of participants.
Therefore, two working memory span tasks were administered to all participants. Based on Baddeley’s model of working memory discussed in section 1.1.2.1, both of these tasks were theoretically verbal working memory tasks where the central executive and the phonological loop were targeted. This is because in both tasks participants were asked to encode, manipulate, and later recall verbal rather than visual-spatial information.

2.1.2.1 Sentence span task.

Waters and Caplan’s (2003) sentence span task, which is a modified version of Daneman and Carpenter’s (1980) reading span task was used. In this task, participants were shown one sentence at a time on a computer screen. All sentences had the form of a subject-cleft (e.g. “It was the child that bit into the fruit). Half of the sentences were grammatical and half were not. Participants were asked to read each sentence on the screen silently to themselves, and make a grammatical judgment for each sentence that required online syntactic analysis. Simultaneously, they were also asked to remember the last word of each sentence. There are five sets of sentences at each span level, with two to eight sentences in each set depending on the span level. After reading all sentences in a set and making grammatical judgments on each of them, participants were then prompted by a blank screen to recall the last word of each sentence in the set in the order they had seen them. The number of sentences in each set (i.e., the span level), and hence the number of words the participants needed to remember, increased as the task progressed.

A participant’s score was the highest span level at which she or he could recall in order all the last words for at least three out of five sets of sentences (regardless of performance on lower span levels). In accordance with the test protocol, an additional 0.5
was added to the score if she or he could recall all the last words for two out of five sets of sentences at any higher span level. This 0.5 addition was only awarded once even if the participant was able to recall two out of five sets of sentences at more than one higher span level. Each participant was tested up to span size six regardless of performance on lower span sizes. The following is an example of how the task was administered and scored. If a participant successfully recalled all the last words for five out of five sets of sentences at span size two, three out of five sets at span size three, two out of five sets at span size four, and zero out of five sets at span five and six, then his or her span score would be 3.5.

The sentence span task was chosen from among other working memory span measures, such as backward digit span and subtract-two span, because various studies have shown that it has more desirable psychometric properties compared to other measures. Summarizing the results from different studies, Waters and Caplan (2003) stated that the sentence span task achieved a fairly high standard in terms of internal consistency as measured by split-half reliability, test-retest reliability, and stability of subject classification (i.e., whether the classification of a particular participant into a working memory group remains unchanged over time and across working memory tasks). They also conducted a study themselves to explore the psychometric properties of various working memory tasks, including the sentence span task, and found similar results. Specifically, while internal consistency was found to be adequate across all the measures examined, the sentence span task achieved the highest test-retest reliability. In fact, according to Waters and Caplan, it was the only measure examined that met the 0.70 criterion that some researchers have argued to be the minimum reliability requirement.
Table 2.2 shows the sentence span score of the participants arranged by age groups. The mean score for the younger and older group are 3.65 ($SD = 1.51$, range = 2 to 8) and 2.05 ($SD = 0.60$, range = 1.5 to 3.5) respectively. An independent sample t-test was conducted to compare the mean scores of the two groups. The difference between the two means was found to be significant, $t(38) = 4.389, p < 0.001$.

2.1.2.2 Counting span task.

The counting span task chosen was adopted from Engle, Tuholski, Laughlin, and Conway (1999). The task was an adapted version of the counting span task originally designed by Case, Kurland, and Goldberg (1982) for children. In this task, participants were presented with one display of different shapes at a time. Each display consisted of three to nine dark blue circles, one to nine dark blue squares, and one to five light blue circles that were randomly arranged. Participants were then instructed to ignore the dark blue square and light blue circles, but to count the number of dark blue circles on the display. They were asked to count out loud and point to each dark blue circle as they counted. Once they had finished counting one display, they were asked to repeat the total, so that the examiner knew that they finished. The examiner would then immediately present the next display, and participants were asked to commence counting as soon as the new display came onto the screen. There are three sets of displays at each span level, with one to eight displays in each set depending on the span level. Like the sentence span task, after counting a number of displays, participants would be prompted by a blank screen to recall the total number of dark circles on each display they had just counted since the last recall in the correct order. The number of displays in each set (i.e., the span level) increased as the task progressed.
A participant’s score was the highest span level at which she or he could recall all the total number of dark blue circles for at least two out of three sets of displays (regardless of performance on lower span levels). An additional 0.3 was added to the score if she or he could recall all the totals for one out of three sets of displays at any higher span level. Similar to the sentence span task, this 0.3 addition was only awarded once even if the participant was able to recall one out of three sets of displays at more than one higher span level. The task progressed until the participant could not recall any of the totals in all three sets of displays at that particular span level. The following is an example of how the task was administered and scored. If a participant successfully recalled all the totals for three out of three sets of displays at span size one and two, two out of three sets at span size three, one out of three sets at span size four, and zero out of five sets at span five, then his or her span score would be 3.3.

Similar to the sentence span task, the counting span task was chosen from among other working memory measures because it possesses a high standard of psychometric properties. Conway, Kane, Bunting, Hambrick, Wilhelm, and Engle (2005) conducted a study that investigated the psychometric properties of various working memory measures, including the counting span task. In their review of the current literature, they stated that in terms of internal consistency, it has been found in one study that the counting span task achieved a coefficient alpha of 0.77. They also showed that scores of the counting span task remained stable across time, and that construct validity of the task was highly satisfactory.

Table 2.1 also shows the counting span score of the participants arranged by groups. The mean score for the younger and older group are 5.15 ($SD = 1.51$, range = 2.3 to 8) and 3.60 ($SD = 1.09$, range = 2.3 to 7) respectively. An independent sample t-test was conducted
to compare the mean scores of the two groups. The difference between the two means was found to be significant, $t(38) = 3.795, p = 0.001$.

Table 2.2 Working memory span scores

<table>
<thead>
<tr>
<th>Subject number</th>
<th>Sentence Counting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sentence</td>
</tr>
<tr>
<td></td>
<td>Young</td>
</tr>
<tr>
<td>1</td>
<td>6.5</td>
</tr>
<tr>
<td>2</td>
<td>4.5</td>
</tr>
<tr>
<td>3</td>
<td>4.5</td>
</tr>
<tr>
<td>4</td>
<td>3.5</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>2.5</td>
</tr>
<tr>
<td>7</td>
<td>2.5</td>
</tr>
<tr>
<td>8</td>
<td>2.5</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>2.5</td>
</tr>
<tr>
<td>11</td>
<td>3.5</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>2.5</td>
</tr>
<tr>
<td>14</td>
<td>4.5</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>16</td>
<td>2.5</td>
</tr>
<tr>
<td>17</td>
<td>4</td>
</tr>
<tr>
<td>18</td>
<td>4</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>4</td>
</tr>
</tbody>
</table>

Mean 3.65 2.05 5.15 3.6
SD 1.51 0.6 1.51 1.09

2.1.3 Lexical knowledge

To measure the participants' lexical knowledge, the North American Adult Reading Test—short form (NAART35) was administered. The long form of the test (Blair and Spreen, 1989, Spreen and Strauss, 1991) was a modification of the National Adult Reading
Test (NART) developed and standardized in England (Nelson, 1982). The NAART35 is essentially a standardized single word reading task. It consists of 35 English words with irregular and unpredictable pronunciations. Participants were asked to read each word out loud, and their raw score was the number of correctly pronounced words. The correct pronunciation of the words was based on the International Phonetic Association (IPA) symbols provided in the Canadian Oxford Dictionary (Barber, 2004). The raw score was then converted into a z-score corrected for age based on a statistical equation provided by Uttl (2002).

Uttl (2002) developed the NAART35 from its long form counterpart, using various psychometric analyses. He excluded items in the long form that had relatively low inter-rater reliability (i.e., a Cohen’s kappa lower than 0.75) or weak power in discriminating examinees who performed well vs. poorly on the test. The items left after the analyses, which had good inter-rater reliability, high discriminatory indices and internal consistencies, form the word list of the NAART35 (See Appendix C for the list of words used in the NAART35). The reliability of the NAARTS 35 scores was reported to be 0.92, using the Cronbach’s alpha. Scores of the NAART35 were also found to have a positive correlation of 0.76 with the scores of the WAIS-R Vocabulary.

Table 2.3 shows the raw score and z-score of the participants by age groups. The mean raw score for the younger and the older age group was 20.75 ($SD = 4.52$, range = 13 to 30) and 24.6 ($SD = 6.41$, range = 13 to 34) respectively. The mean z-score for the younger and older group are -0.04 ($SD = 0.64$, range = -1.12 to 1.19) and 0.11 ($SD = 0.89$, range = -1.53 to 1.39) respectively. An independent sample t-test was conducted to compare the mean raw score and the mean z-score of the two groups. A significant difference was found
between the mean raw score of the age groups, $t(38) = -2.195, p = 0.034$, meaning that the older participants had significantly richer lexical knowledge relative to the younger participants. However no difference was found between the mean z-score of the age groups, $t(38) = -0.635, p = 0.529$, which suggested that their lexical knowledge was comparable to that of their same-age peers.

Table 2.3 NAART35 raw and z-score

<table>
<thead>
<tr>
<th>Subject number</th>
<th>Young raw</th>
<th>Young z-score</th>
<th>Old raw</th>
<th>Old z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30</td>
<td>1.19</td>
<td>33</td>
<td>1.2</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>-0.14</td>
<td>25</td>
<td>0.16</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>-1.12</td>
<td>19</td>
<td>-0.65</td>
</tr>
<tr>
<td>4</td>
<td>17</td>
<td>-0.57</td>
<td>22</td>
<td>-0.24</td>
</tr>
<tr>
<td>5</td>
<td>18</td>
<td>-0.48</td>
<td>24</td>
<td>0.07</td>
</tr>
<tr>
<td>6</td>
<td>24</td>
<td>0.48</td>
<td>13</td>
<td>-1.53</td>
</tr>
<tr>
<td>7</td>
<td>24</td>
<td>0.44</td>
<td>15</td>
<td>-1.31</td>
</tr>
<tr>
<td>8</td>
<td>14</td>
<td>-0.97</td>
<td>19</td>
<td>-0.64</td>
</tr>
<tr>
<td>9</td>
<td>17</td>
<td>-0.58</td>
<td>22</td>
<td>-0.27</td>
</tr>
<tr>
<td>10</td>
<td>19</td>
<td>-0.26</td>
<td>23</td>
<td>-0.09</td>
</tr>
<tr>
<td>11</td>
<td>18</td>
<td>-0.45</td>
<td>16</td>
<td>-1.06</td>
</tr>
<tr>
<td>12</td>
<td>26</td>
<td>0.71</td>
<td>34</td>
<td>1.39</td>
</tr>
<tr>
<td>13</td>
<td>17</td>
<td>-0.61</td>
<td>26</td>
<td>0.35</td>
</tr>
<tr>
<td>14</td>
<td>21</td>
<td>0.03</td>
<td>24</td>
<td>0.05</td>
</tr>
<tr>
<td>15</td>
<td>20</td>
<td>-0.18</td>
<td>23</td>
<td>-0.12</td>
</tr>
<tr>
<td>16</td>
<td>27</td>
<td>0.86</td>
<td>33</td>
<td>1.2</td>
</tr>
<tr>
<td>17</td>
<td>22</td>
<td>0.17</td>
<td>32</td>
<td>1.17</td>
</tr>
<tr>
<td>18</td>
<td>24</td>
<td>0.43</td>
<td>33</td>
<td>1.32</td>
</tr>
<tr>
<td>19</td>
<td>26</td>
<td>0.66</td>
<td>31</td>
<td>1.02</td>
</tr>
<tr>
<td>20</td>
<td>18</td>
<td>-0.45</td>
<td>25</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Mean 20.75  -0.042  24.6  0.11
SD 4.52  0.64  6.41  0.89
2.2 Experimental Materials and Tasks

Participants were asked to complete two narrative production tasks, each requiring them to tell a story based on the visual stimuli presented. The first story, which depicted a cowboy’s misadventure, was a modified version of one appearing in Ombredane’s “The Miracle” (1951) as cited and used in Joanette and Goulet’s (1990) study. The second story, which illustrated a fisherman’s mishap, was designed specifically for this study, with the intention of creating a story that was comparable to the Cowboy’s Story. In terms of the number of main characters and setting, both stories involved two male main characters, an animal character, and a comparable number of objects in the setting (7 for the Cowboy’s Story and 8 for the Fisherman’s Tale). In terms of theme and development, both stories portrayed a trick being played on one of the male characters by the other male character.

The two stories were each represented by a picture sequence and a silent video. The picture sequence of the Cowboy’s Story consists of eight black-and-white hand-drawn frames. The first three frames of each story illustrate the setting, which are followed by four frames of complication. The last frame in each picture sequence portrays the resolution. The video of the Cowboy’s Story was made by a group of students from the School of Audiology and Speech Sciences, University of British Columbia, for a class assignment (Fearon, Shaver & Reich, 2008). The video is one minute and fifty-two seconds long. It is in colour but silent, and it has the same content as its picture sequence counterpart.

For the Fisherman’s Tale, the video was made before the picture sequence was drawn. The video is one minute and sixteen seconds long. Like the video of the Cowboy’s Story, it is in colour and silent. The picture sequence was then hand-drawn based on the
video such that the two contained equivalent content. It was also drawn in the same style and had the same macrostructure (i.e., the first three frames illustrating the setting, the next four the complication, and the last the resolution) as the picture sequence of the Cowboy's Story. The picture sequence of the Cowboy’s Story can be found in Joanette and Goulet (1990, p. 133) while the Fisherman’s Tale can be found in Appendix D.

Because the Fisherman’s Tale was composed for the purpose of this study, and hence, its video and picture sequence had never been used as visual stimuli for any narrative tasks, a pilot study was conducted to ensure that the video and the picture sequence would elicit narratives that resembled the intended story. Nine individuals, who were all graduate students at the School of Audiology and Speech Sciences, University of British Columbia, were recruited in the pilot study. Five produced a narrative from the video and four from the picture sequence. All narratives produced in both conditions contained contents that closely resembled the intended story.

2.3 Procedure

Participants in both the younger and older group each produced one narrative for each story. Half of the participants in each group used a picture sequence of the Cowboy’s Story and a video of the Fisherman’s Tale as visual stimuli. The other half of the participants in each group used a video of the Cowboy’s Story and a picture sequence of the Fisherman’s Tale. The order in which the picture sequence and the video condition were presented for each participant was counterbalanced. This yielded four subgroups of participants: (1) Young, cowboy picture, fisherman video (Y-CP-FV); (2) Young, cowboy video, fisherman picture (Y-CV-FP); (3) Old, cowboy picture, fisherman video (O-CP-FV); and (4) Old,
cowboy video, fisherman picture O-CV-FP). The narrative production tasks took place in a quiet room. In the picture sequence condition, participants were instructed to look at the eight picture frames of the corresponding story and formulate a narrative based on the picture sequence. Participants were encouraged to take as much time as they needed. They were also allowed to look at the picture sequence while they were producing the narrative. In the video condition, participants were seated in front of a computer where they were shown the video of the corresponding story and were asked to formulate a story based on the video. As discussed in section 1.5, participants were not allowed to look at the video while they produced the narrative, as this would implicitly impose a time constraint on the narrative production of the participants. However, participants were encouraged to watch the video as many times as they needed before producing the narrative. In both the picture sequence and the video condition, participants were encouraged to produce a narrative that was as informative and descriptive of what was happening in the stimulus as possible. The examiner did not answer any questions asked by the participants while they were producing the narratives. All narratives were recorded using a digital recorder Marantz Professional Solid State Recorder PMD661.

2.4 Analysis

2.4.1 Transcription and reliability

Because each participant produced two stories, a total of 80 narratives were collected. All recorded stories were transcribed and coded using the computer program CHAT of the CHILDES project (MacWhinney, 2000). In terms of unit segment, each utterance consists of one T-unit, which, according to Paul (2007), is essentially a main clause and all of its
subordinate clauses. As a result, two clauses joined by a subordinating conjunction such as "because", "since", and “although” were transcribed as one utterance, while two conjoined clauses joined by a coordinating conjunction such as “and” and “but” were transcribed as two utterances. For the conjunction “so”, when it functioned syntactically as a coordinating conjunction, the two conjoined clauses were transcribed as two separate utterances. When it functioned as a subordinating conjunction with the meaning of “so that”, the two clauses were transcribed as one utterance. Discourse behaviours such as repetition, reformulation and pauses were coded as specified by the CHAT coding manual (See Appendix E for a sample transcript and the list of codes used in transcription). Ten percent of the collected narratives were transcribed by a second transcriber. Inter-rater reliability of the transcription was 95% in terms of utterance and 97% in terms of words. In cases of discrepancy, a third transcriber was consulted such that an agreement between at least two out of three transcribers was achieved on each discrepancy. All transcripts were verified by a second listener.

2.4.2 Dependent variables

2.4.2.1 Type-token ratio and verb-noun ratio.

TTR and verb-noun ratio were calculated as lexical measures. TTR was calculated for each transcript, using the CLAN program of the CHILDES project (MacWhinney, 2000). The program identifies the total number of words and the total number of different word roots for each transcript, from which TTR is calculated. Verb-noun ratio for each transcript was calculated by manually counting the number of nouns and verbs in each transcript. For nouns, common nouns, proper nouns (such as names), and pronouns were included in the counting. Compound nouns (e.g., “cowboy hat” and “toy horse”) were only counted as one
noun, because they semantically and syntactically function as one unit. For verbs, only main verbs, as opposed to auxiliary verbs, such as the auxiliary copula “be” and modal verbs, were counted. In the case of gerunds, despite the fact that their syntactic function resembles that of nouns, they were still considered to be verbs. On the other hand, when the participle or perfect form of a verb was used as an adjective (e.g. “the sleeping cowboy” and “the broken fishing rod”), it was excluded from the verb count.

2.4.2.2 Cohesion and coherence.

To measure cohesion and coherence, referencing, non-progression, contradiction, and irrelevance errors, based on the definitions provided by Joanette and Goulet (1990), were coded in the CHAT format on the transcripts (MacWhinney, 2000; see Appendix F for the detailed list and description of the error codes). The CLAN program was then used to count the total number of each type of errors for each narrative. Cohesion and coherence errors were considered separately in the analysis. That is, each participant received one score for cohesion errors and one score for coherence errors for each narrative produced. The cohesion error score equalled the total number of reference errors in a narrative, while the coherence error score equalled the total number of non-progression, contradiction and irrelevance errors combined in a narrative.

Cohesion and coherence error reliability coding was conducted separately on 10% of the transcripts by a second rater. The inter-rater reliability of the cohesion and coherence error coding was 64%. The relatively low inter-rater reliability was speculated to be due to the small total number of errors in the transcripts examined. Nevertheless, discrepancies between the coding of the two raters were discussed. A second Inter-rater reliability check
was then conducted on an additional 10% of the transcripts. The total number of errors in the transcripts was still found to be small. However, inter-rater reliability improved to 75%.

2.4.2.3 Proposition density.

In terms of content richness, a proposition analysis was conducted for each transcript based on the method proposed by Turner and Greene (1977). Recall that a proposition consists of a relation word concept and the argument word concepts it takes. Relation word concepts can be a predicate (e.g., verbs), a modifier (e.g., adjectives and adverbs), or a connective (e.g., coordinating and subordinating conjunctions), while argument word concepts were usually noun phrases, main clauses and subordinating clauses. Propositions could also be simple or complex. In the former, no arguments of the proposition corresponded to another proposition (Joanette & Goulet, 1990). On the other hand, if a proposition contained an argument that corresponded to another proposition, then the proposition was considered to be complex. Propositions were extracted from each utterance of each transcript such that a connected propositional text base was constructed for each narrative (See Appendix G for a sample proposition analysis based on the sample transcript in Appendix E.

Proposition analyses were conducted on 10% of the transcripts by a second rater. Inter-rater reliability in terms of the total number of propositions was found to be 89%. Discrepancies in the analyses between the two raters were discussed and resolved. The P-density of each of the 80 transcripts was then calculated by dividing the total number of propositions by the total number of words produced in each narrative (Kemper & Sumner, 2001).
2.4.3 Core proposition analysis

As discussed earlier, the Cowboy’s Story of the current study was adopted from Joanette and Goulet’s (1990) study where they included narrative discourse of individuals with and without right-hemisphere damage. In their assessment of content richness of the narratives in their study, a list of 32 core propositions, which were the expected informative units that formed the gist of the Cowboy’s Story produced by at least 20% of their participants, was identified (see Appendix H for the list of core propositions). In order to compare the narratives of the participants of the current study with those of the normal healthy control participants in Joanette and Goulet’s study ($N = 20$; mean age = approximately 57; mean years of education = approximately 8; $SD$ and range not provided for either measure), the Cowboy’s Story narrative of each participant, whether elicited in the picture sequence or video condition, was evaluated with respect to Joanette and Goulet’s list of core propositions. Specifically, the percentage of core propositions produced was calculated for each Cowboy’s Story narrative. The mean percentage of core propositions produced in the current study was then compared to that of the normal healthy control group in Joanette and Goulet’s study. The core proposition analysis was only conducted on the Cowboy’s Story narratives because a list of core propositions has yet to be developed for the Fisherman’s tale.

2.4.4 Statistical analysis

Statistical analysis of the data, using the computer statistics program SPSS was conducted in order to determine whether differences in performance found between different subgroups of participants reached the level of significance. Five separate ANOVAS with
repeated measures were conducted, one for each dependent variable, to identify any main and interaction effects the independent variables (age, task, order and combination of story and task) had on the dependent measures (TTR, verb-noun ratio, cohesion error, coherence error and P-density). Age was a between-subjects factor while task (picture vs. video) was a within-subjects factor. The order of task presentation (i.e., picture first vs. video first) and the combination of story and task (i.e., CP-FV vs. CV-FP) were also included as between-subject factors in the univariate tests. While the variable order is quite self-explanatory, the variable combination requires some explanation. Recall that the Fisherman’s Tale was designed to match the Cowboy’s Story in terms of the number of main characters, objects, and macrostructure in order to minimize any effects of the stories on the dependent measures. However, if the two stories did indeed behave differently regardless of the attempt to have the two stories matched, then how the stories were grouped with the task conditions (i.e., CP-FV vs. CV-FP) would be anticipated to affect the resulting narratives differently. Therefore the inclusion of combination as between-subjects factor in the analyses was to verify that the two stories were behaving similarly and did not differentially affect the dependent measures.

A univariate test, as opposed to a multivariate test (e.g., a 2 x 2 MANOVA), was chosen because a multivariate test requires all the dependent measures to be correlated. While the five dependent variables in this study were all quantitative measurement of narratives, their correlation appeared conceptually neither straightforward nor significant. Therefore, univariate testing was used instead. On the other hand, because five separate univariate analyses was conducted, the alpha level of each analysis needed to be adjusted by the Bonferroni correction in order to guard against Type I errors, yielding a alpha level of 0.01 for each univariate analysis.
In addition to the ANOVAs conducted on each of the dependent measures, an
ANOVA was also conducted to identify any main and interaction effects age and task had on
the percentage of core propositions produced in the Cowboy's Story narrative. Such analysis
was conducted for exploratory purposes and was motivated by the fact that content richness
measured in core or main propositions produced, in addition to proposition density, had also
been shown to be affected by age and task condition (see Chapter 1 for specific research
studies). Given that the sample size decreased by half in this analysis because only the
Cowboy's Story narratives were analysed, it was not considered as part of the series of
ANOVA for the dependent measures. Therefore, the alpha level was not corrected by the
Bonferroni adjustment and was set at 0.05 for this analysis.
3. Results

Following the procedures outlined in Chapter 2, a total of 80 narrative samples were collected from the four groups of participants and analysed accordingly. In this chapter, the results of each dependent measure are compared across age groups and task conditions. To determine the statistical significance of any difference found between the age groups and task conditions (i.e., the two independent variables of primary interest), results of the univariate test for each dependent measure are also presented. These results are followed by an investigation of order and combination effects on the dependent measures and the results of the core proposition analysis for the Cowboy’s story narratives. This chapter is concluded with a summary of all the findings.

3.1 Summary of Data

Table 3.1 summarizes the mean scores and the standard deviation of each dependent variable by story, age group and task condition. In order to find out whether the two independent variable of interest, namely, age and task conditions had an effect on the dependent measures, results of each dependent variable were collapsed across the two stories. This yielded the data presented in Table 3.2. In the following section, results of each dependent variable are compared across age groups and task conditions based on the results shown in Table 3.2. Results of the ANOVA for each dependent variable are also presented. Recall that because multiple univariate tests were conducted, the alpha level of each ANOVA was corrected to 0.01 by a Bonferroni adjustment.
Table 3.1 Summary of data arranged by story, task condition and age group

<table>
<thead>
<tr>
<th></th>
<th>Cowboy’s Story</th>
<th>Fisherman’s Tale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Picture Sequence</td>
<td>Video</td>
</tr>
<tr>
<td></td>
<td>Young</td>
<td>Old</td>
</tr>
<tr>
<td><strong>TTR</strong></td>
<td>Mean</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.06</td>
</tr>
<tr>
<td><strong>VN</strong></td>
<td>Mean</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.08</td>
</tr>
<tr>
<td><strong>Cohesion errors</strong></td>
<td>Mean</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.67</td>
</tr>
<tr>
<td><strong>Coherence errors</strong></td>
<td>Mean</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.32</td>
</tr>
<tr>
<td><strong>PD</strong></td>
<td>Mean</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.04</td>
</tr>
</tbody>
</table>
Table 3.2 Summary of data arranged by age group and task condition

<table>
<thead>
<tr>
<th></th>
<th>Young</th>
<th>Old</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TTR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture</td>
<td>0.481</td>
<td>0.495</td>
<td>0.488</td>
<td>0.064</td>
</tr>
<tr>
<td>Video</td>
<td>0.458</td>
<td>0.459</td>
<td>0.458</td>
<td>0.061</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>0.469</td>
<td>0.477</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>0.061</td>
<td>0.068</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>VN</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture</td>
<td>0.697</td>
<td>0.678</td>
<td>0.687</td>
<td>0.099</td>
</tr>
<tr>
<td>Video</td>
<td>0.662</td>
<td>0.673</td>
<td>0.667</td>
<td>0.105</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>0.679</td>
<td>0.675</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>0.094</td>
<td>0.110</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cohesion error</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture</td>
<td>0.5</td>
<td>1</td>
<td>0.75</td>
<td>0.840</td>
</tr>
<tr>
<td>Video</td>
<td>0.55</td>
<td>0.7</td>
<td>0.625</td>
<td>0.925</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>0.525</td>
<td>0.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>0.640</td>
<td>1.051</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Coherence error</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture</td>
<td>0.24</td>
<td>0.6</td>
<td>0.425</td>
<td>1.059</td>
</tr>
<tr>
<td>Video</td>
<td>0.45</td>
<td>0.6</td>
<td>0.525</td>
<td>0.784</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>0.35</td>
<td>0.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>0.622</td>
<td>1.150</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>P-density</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture</td>
<td>0.400</td>
<td>0.417</td>
<td>0.409</td>
<td>0.041</td>
</tr>
<tr>
<td>Video</td>
<td>0.390</td>
<td>0.403</td>
<td>0.397</td>
<td>0.050</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>0.395</td>
<td>0.410</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>0.040</td>
<td>0.050</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.2 Analyses of the Dependent Variables

3.2.1 Type-token ratio

Table 3.2 shows that type-token ratio (TTR) appeared to be higher in narratives produced by the older age group ($M = 0.469, SD = 0.061$ for the younger age group; $M = 0.477, SD = 0.068$ for the older age group). TTR also appeared to be higher in narratives elicited in the picture condition ($M = 0.488, SD = 0.064$) than those elicited in the video condition ($M = 0.458, SD = 0.061$). While the difference between the age groups failed to
reach a statistical significance, $F(1, 32) = 0.195, p = 0.662, \eta^2_p = 0.006$, the difference between the task conditions was found to be significant, $F(1, 32) = 10.293, p = 0.003, \eta^2_p = 0.243$. No interaction effect of age and task was found, $F(1, 32) = 0.426, p = 0.519, \eta^2_p = 0.013$. Figure 3.1 and 3.2 summarize the results of type-token ratio by age group and task condition respectively.

![Fig 3.1 TTR by age](image1)

![Fig 3.2 TTR by task](image2)

The fact that a main effect of task condition on type-token ratio was found called for further investigation. Specifically, TTR measure has been shown to be sensitive to the size of the given language sample in terms of total number of words (TNW) in research on both child language and acquired language disorders (Wrigth, Silverman, & Newhoff, 2003).
Larger language samples tend to yield a lower TTR. Therefore, in order to determine whether the main effect of task on TTR was at least partially due to a difference in the TNW across the task conditions, the mean TNW in narratives elicited from the picture sequence condition was compared to those elicited from the video condition.

Figure 3.3 shows the mean TNW of the narratives elicited from the two task conditions (\(M = 187.08, SD = 90.32\) for the picture sequence condition; \(M = 243.85, SD = 95.21\) for the video condition). A paired sample t-test was then conducted to compare the means of the 2 task groups and a significant difference was found, \(t(39) = -4.200, p = 0.000\). In light of such difference, a post-hoc ANCOVA with repeated measures was carried out on the dependent variable TTR, in which TNW was included as a covariate in the analysis. Results of the ANCOVA showed that once TNW was included as a covariate, the main effect of task condition on TTR was no longer significant, \(F(1, 38) = 0.300, p = 0.589, \eta^2_p = 0.008\).
Instead TNW accounted for most of the difference found between the mean TTR of the two task conditions, $t = -4.730, p < 0.001$.

### 3.2.2 Verb-noun ratio

Verb-noun ratio was found to be marginally higher in the younger age group ($M = 0.679$, $SD = 0.094$) than the older age group ($M = 0.675$, $SD = 0.110$). Comparing the results across task conditions, it was found that verb-noun ratio of narratives elicited from the picture sequence condition ($M = 0.687$, $SD = 0.099$) was on average higher than that of narratives elicited from the video condition ($M = 0.667$, $SD = 0.105$). However, results of the ANOVA showed that neither the effect of age, $F(1, 32) = 0.027, p = 0.871, \eta^2_p = 0.001$, nor task, $F(1, 32) = 0.864, p = 0.360, \eta^2_p = 0.026$ was significant. There was also no significant interaction effect of age and task, $F(1, 32) = 0.513, p = 0.479, \eta^2_p = 0.016$. Figure 3.4 and 3.5 summarize the results of verb-noun ratio by age group and task condition respectively.

![Fig. 3.4 Verb-noun ratio by age](image1)

![Fig. 3.5 Verb-noun ratio by task](image2)
3.2.3 Cohesion

As shown in Table 3.2, participants in the older age group on average committed more cohesion errors ($M = 0.85$, $SD = 1.051$) than participants in the younger age group ($M = 0.525$, $SD = 0.640$). Narratives elicited from the picture sequence condition ($M = 0.75$, $SD = 0.840$) also contained more cohesion errors on average than those elicited from the video condition ($M = 0.625$, $SD = 0.925$). However, similar to verb-noun ratio, neither age, $F(1, 32) = 2.991, p = 0.093, \eta^2_p = 0.085$ nor task, $F(1, 32) = 0.407, p = 0.528, \eta^2_p = 0.013$ had a significant main effect on cohesion error. Also, no significant interaction effect of age and task was detected, $F(1, 32) = 0.797, p = 0.379, \eta^2_p = 0.024$. Figure 3.6 and 3.7 summarize the results of cohesion error by age group and task condition respectively.

Fig. 3.6 Cohesion error by age

Fig. 3.7 Cohesion error by task
3.2.4 Coherence

Participants in the older age group on average committed more coherence errors ($M = 0.6$, $SD = 1.150$) than participants in the younger age group ($M = 0.35$, $SD = 0.622$). The mean coherence error of the narratives in video condition ($M = 0.525$, $SD = 0.784$) was also found to be higher than that of the narratives in the picture sequence condition ($M = 0.425$, $SD = 1.059$). Results of the ANOVA revealed that both age, $F(1, 32) = 1.379, p = 0.249, \eta^2_p = 0.041$ and task, $F(1, 32) = 0.224, p = 0.639, \eta^2_p = 0.007$ did not have a significant main effect on coherence error. Also, no significant interaction effect of age and task was noted, $F(1, 32) = 0.224, p = 0.639, \eta^2_p = 0.007$. Figure 3.8 and 3.9 summarize the results of coherence error by age group and task condition respectively.

![Fig. 3.8 Coherence error by age](image1)

![Fig 3.9 Coherence error by task](image2)
3.2.5 Content richness

Table 3.1 shows that proposition density was higher for the older age group \( (M = 0.410, SD = 0.050) \) than the younger age group \( (M = 0.395, SD = 0.040) \). Narratives in the picture sequence condition also had a slightly higher proposition density \( (M = 0.409, SD = 0.041) \) than the narratives in the video condition \( (M = 0.397, SD = 0.050) \). However, results of the univariate test revealed that the difference in the means between the two age groups and between the task conditions was not significant. In other words, similar to cohesion and coherence error, neither age, \( F(1, 32) = 1.505, p = 0.229, \eta^2_p = 0.045 \) nor task condition, \( F(1, 32) = 2.705, p = 0.110, \eta^2_p = 0.078 \) had a main effect on proposition density. There was also no significant interaction effect of age and task on proposition density, \( F(1, 32) = 0.091, p = 0.765, \eta^2_p = 0.003 \). Figure 3.10 and 3.11 summarize the results of proposition density by age groups and task condition respectively.

Fig. 3.10 P-density by age

Fig. 3.11 P-density by task
3.3 Order and Combination Effects

As explained in Chapter 2, order and combination were included as between-subject variables in the univariate test for each of the dependent measure to verify that the order of task presentation (picture first vs. video first) and how the stories were paired with the task conditions (CP-FV vs. CV-FP) did not affect the resulting narratives. Results of the univariate tests showed that order and combination did not have a significant main effect or interaction effect with other independent variables. However, it is worth noting that marginally significant main and interaction effects were noted for four out of five of the dependent measures. Table 3.3 summarizes the marginally significant main and interaction effects found in the univariate tests. Specifically, while marginal order effects were only noted in interaction with other independent variables on two dependent measures, namely, TTR and verb-noun ratio, marginal main or interaction effects of story-task combination on TTR, verb-noun ratio, cohesion and proposition density were noted.

Table 3.3 Marginally significant effects of order and story-task combination

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Dependent Variable</th>
<th>F(1, 32)</th>
<th>p</th>
<th>η²_p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combination</td>
<td>Cohesion error</td>
<td>5.115</td>
<td>0.031</td>
<td>0.138</td>
</tr>
<tr>
<td>Task * order * combination</td>
<td>TTR</td>
<td>5.199</td>
<td>0.029</td>
<td>0.14</td>
</tr>
<tr>
<td>Age * order</td>
<td>Verb-noun ratio</td>
<td>5.355</td>
<td>0.027</td>
<td>0.143</td>
</tr>
<tr>
<td>Age * combination</td>
<td>Verb-noun ratio</td>
<td>6.129</td>
<td>0.019</td>
<td>0.161</td>
</tr>
<tr>
<td>Task * age * combination</td>
<td>Proposition density</td>
<td>5.253</td>
<td>0.029</td>
<td>0.141</td>
</tr>
</tbody>
</table>

These marginally significant effects should be taken with caution. Recall that the alpha level was set at 0.01 by a Bonferroni adjustment. However, the Bonferroni adjustment has been shown to be overly conservative in controlling for Type I error (MacDonald,
Seifert, Lorenzet, Givens & Jaccard, 2002). Effects that are deemed insignificant but have $p$
values that are close to the alpha level could actually be significant. In other words,
correcting the alpha level by the Bonferroni adjustment may increase the likelihood of a Type
II error. Therefore, the findings summarized in Table 3.2 were suggestive of the fact that
there was possibly an effect of order and even more probably combination on at least some of
the dependent measures, despite counterbalancing the order of task presentation and the
attempt to design the Fisherman’s Tale in the fashion to match the Cowboy’s Story.

Because order effect was only observed in interaction with other independent
variables on two of the dependent measures, and given that there was no theoretical or
practical reason to assume that the order of task presentation would influence the
participants’ performance on the dependent measures, order effect was not further explored.
On the other hand, combination effect was observed either as main or interaction effect on
four of the dependent measures. Also, it was likely that differences between the two stories
still existed despite having them matched, since they depicted two different events after all.
Therefore, combination effect was further investigated. Specifically, because the presence of
marginal combination effects suggested that the two stories may be behaving differently, a
post-hoc analysis where the narratives of the two stories were analysed separately was
conducted.

In these analyses, one for the Cowboy’s Story narratives and one for the Fisherman’s
Tale narratives, a series of five univariate tests, one for each dependent variable with alpha
level set at 0.01 by a Bonferroni adjustment, were conducted. The independent variables
were age and task. Because the stories were now analysed separately, the sample size
decreased by half in their respective analysis (i.e., 40 Cowboy’s Story narratives and 40
Fisherman’s Tale narratives). Also because of the two stories being analysed separately, task became a between-subject variable in the univariate tests (i.e., 20 picture sequence narratives vs. 20 video narratives produced by different participants). Results of the separate analyses of the two stories are presented below.

3.3.1 Cowboy’s Story

As in the results of the previous analysis where the two stories were analysed collectively, there was again no significant main effect of age or any interaction effect of age and task on any of the dependent measures. TTR of the narratives elicited from the picture sequence condition (\(M = 0.503, SD = 0.057\)) was again significantly higher than those elicited from the video condition (\(M = 0.439, SD = 0.050\)), \(F(1, 32) = 14.262, p = 0.001, \eta^2_p = 0.308\). Figure 3.12 shows the mean TTR of the cowboy story arranged by task condition.

![Fig 3.12 Cowboy's Story TTR by task](image)

Given the fact that TNW was found to be a significant covariate of TTR in the previous analysis where the two stories were analysed collectively, a *post-hoc* ANCOVA
was conducted. Similar to the previous analysis, the main effect of task on TTR became insignificant once TNW was included as a covariate in the univariate test, $F(1, 31) = 0.381, p = 0.542, \eta^2_p = 0.012$. Instead the effect of TNW on TTR was significant, $F(1, 31) = 33.503, p = 0.000, \eta^2_p = 0.519$.

However, unlike the previous analysis, task was also found to have a main effect on cohesion error. Specifically, the mean number of cohesion errors was significantly higher in the picture sequence condition ($M = 0.95, SD = 0.76$) than in the video condition ($M = 0.40, SD = 0.60$) for the Cowboy’s Story, $F(1, 32) = 7.806, p = 0.009, \eta^2_p = 0.196$. Figure 3.13 summarizes the mean cohesion error of the two task conditions for the Cowboy’s Story.

![Figure 3.13](image-url)
3.3.2 Fisherman's Tale

Results of the ANOVAs on the dependent measures of the Fisherman's Tale narratives were also similar to those of the Cowboy's Story narratives and the analyses of the two stories combined. No significant main effect of age and task was found on any of the dependent measures. There was also no significant interaction effect of age and task on four out of five of the dependent measures, namely, TTR, cohesion error, coherence error and proposition density. However, it is worth noting that unlike the Cowboy's Story, task did not have a significant main effect on TTR. That is, the mean TTR of the narratives elicited from the picture sequence condition ($M = 0.473$, $SD = 0.070$) was not significantly different from that of the narratives elicited from the video condition ($M = 0.478$, $SD = 0.066$), $F(1, 32) = 0.055, p = 0.816, \eta^2_p = 0.002$. Figure 3.14 shows the mean TTR of the narratives in the picture sequence vs. the video condition. To verify the close relationship observed between TNW and TTR previously, the effect of task on TNW was examined. Because task did not have a main effect on TTR, one would not expect a significant difference in TNW across task conditions as well. A post-hoc univariate test was conducted and the result showed that task indeed did not have a main effect on TNW, $F(1, 32) = 0.068, p = 0.796, \eta^2_p = 0.002$. Given the observation that task has a significant main effect on TTR for the narratives of the Cowboy's Story but not for those of the Fisherman's Tale, one may conclude that the former largely contributed to the significant main effect of task found on TTR in the univariate test reported in section 3.2.1, where the two stories were analysed collectively.
While the interaction of age and task was not found to have any effects on any of the dependent measures for the Cowboy’s Story narratives, a marginally significant interaction effect of age and task was found on verb-noun ratio of the Fisherman’s Tale narratives, $F(1, 32) = 6.384, p = 0.017, \eta^2_p = 0.166$. Individuals in the younger age group on average had a higher verb-noun ratio in the picture sequence condition ($M = 0.705, SD = 0.097$) than the video condition ($M = 0.640, SD = 0.060$). However, the opposite pattern was observed for the older age group, where participants on average produced narratives with a higher verb-noun ratio in the video condition ($M = 0.720, SD = 0.116$) than in the picture sequence condition ($M = 0.640, SD = 0.092$). Figure 3.15 shows the verb-noun ratio for the Fisherman’s Tale arranged by age and task condition.
3.3.3 Summary

Although the separate analyses of the narratives of the two stories and original analysis where the two stories were analysed collectively yielded fairly comparable results, age and task conditions were nevertheless found to have slightly different effects on some of the dependent measures. For example, the significant effect of task on TTR was only observed in the Cowboy’s Story but not the Fisherman’s Tale when the two stories were analysed separately. Also, a significant main effect of task was found on cohesion error for the Cowboy’s Story but not the Fisherman’s Tale. On the other hand, no interaction effect of age and task was found on any dependent measures for the Cowboy’s Story narratives, while a significant interaction of age and task on verb-noun ratio was noted for the Fisherman’s Tale. These observations all suggested that the Cowboy’s Story and the Fisherman’s Tale were not equivalent, despite the original effort to match the two stories. The comparison of
the results yielded from the collective and separate analyses of the two stories is summarized in Table 3.4.

Table 3.4 Comparison of the collective vs. the separate analyses of the two stories

<table>
<thead>
<tr>
<th>Significant effects</th>
<th>Collective</th>
<th>Cowboy</th>
<th>Fisherman</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>task</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>interaction</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Verb-noun ratio</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>task</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>interaction</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Cohesion error</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>task</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>interaction</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Coherence error</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>task</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>interaction</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>P-density</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>task</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>interaction</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

3.4 Core Proposition Analysis of Cowboy’s Story

Based on the procedure described in Chapter 2, the percentage of core propositions produced was calculated for each Cowboy’s Story narrative. The mean percentage of core propositions produced for the Cowboy’s Story in the current study was then compared to that of the normal healthy control group in Joanette and Goulet’s (1990) study. Results of the comparison showed that participants in the current study produced considerably more core propositions of the Cowboy’s Story ($M = 69.45\%, SD = 9.97$, range = 50% to 87.5%) than
the normal healthy control group of Joanette and Goulet's (1990) study (\(M = 51.25\%\), range = 28.13\% to 75 \%, SD not provided).

![Fig 3.16 Mean percentage of core proposition by age](image1)

![Fig 3.17 Mean percentage of core proposition by task](image2)

The mean percentage of core propositions produced for the Cowboy's Story in the present study was also examined across the two different age groups and task conditions for exploratory purposes. The mean percentage of core propositions produced by the older age group (\(M = 69.69\%, SD = 10.14\)) was found to be similar to that produced by the younger age group (\(M = 69.22\%, SD = 10.05\)). The mean percentage of core propositions produced in the video condition was found to be higher (\(M = 72.81\%, SD = 9.68\)) than that produced in the picture sequence condition (\(M = 66.09\%, SD = 9.31\)). A univariate test was conducted with alpha level set at 0.05. It revealed that the difference between the means of the two age
groups was not significant $F(1, 36) = 0.023, p = 0.880, \eta^2_p = 0.001$. On the other hand, the difference between the means of the two task conditions was significant, $F(1, 36) = 4.775, p = 0.035, \eta^2_p = 0.117$. No significant interaction effect of age and task condition was found, $F(1, 36) = 0.209, p = 0.650, \eta^2_p = 0.006$. Figure 3.16 and Figure 3.17 shows the mean percentage of core propositions produced arranged by age groups and by task conditions respectively.

3.5 Summary of Findings

1. Age did not appear to have any main effects on any of the dependent measures (i.e., TTR, verb-noun ratio, cohesion error, coherence error, and proposition density).

2. The use of different visual stimuli, namely, picture sequence vs. video, did not appear to have any significant effect on any of the dependent measures except for TTR.

3. The significant main effect of task on TTR showed that such effect was largely accounted for by the significant difference between the mean total numbers of words (TNW) of narratives elicited from the two task conditions.

4. There was also no significant interaction effect of age and task on the dependent variables.

5. The marginally significant combination effects on the dependent measures and the fact that analysing the Cowboy’s Story and the Fisherman’s Tale separately yielded slightly discrepant results suggested that the two stories were not equivalent, and thus may impact the dependent measures differentially, despite the initial assumption that the two stories were matched.
6. The mean percentage of core proposition produced for the Cowboy Story in the current study was higher compared to that of the normal healthy control group of Joanette and Goulet's (1990) study.

7. A main effect of task, but not age, was found on the mean core propositions produced for the Cowboy's Story.
4. Discussion

This study was designed to explore the effects of aging and the use of different visual stimuli on narrative production by normal healthy adults in terms of lexical diversity, verb-noun ratio, cohesion and coherence, and content richness. Twenty younger and 20 older adults without a history of cognitive or linguistic deficits participated in this study, and each of them was asked to produce two narratives, one based on a picture sequence stimulus and another on a video stimulus. The narratives were transcribed and type-token ratio, verb-noun ratio, cohesion errors, coherence errors and proposition density were measured for each narrative. Statistical analyses were conducted to compare the results of each of the above dependent measures across the two age groups and the two task conditions (picture sequence vs. video). This chapter reviews and discusses the findings in relation to the research hypotheses proposed in Chapter 1. Other important findings of the study are also discussed. This is followed by a discussion of the limitations of this study, the clinical implications of the findings, and directions for future research.

4.1 Research Hypotheses

4.1.1 The effects of age on lexical diversity, verb-noun ratio, and coherence

The first hypothesis of the current study was that age would not have an effect on lexical diversity, verb-noun ratio or coherence in narrative production. The results largely supported this hypothesis.
4.1.1.1 Age and lexical diversity.

Lexical diversity was measured in type-token ratio (TTR), which was calculated by dividing the number of different words by the total number of words in a given piece of narrative. Findings of the current study showed that while the mean TTR was slightly higher for the older age group than the younger age group, this difference did not reach statistical significance. This finding is in line with the results of Nef and Hupet’s (1982) study and Kemper and Sumner’s (2001) study, where the effect of age on lexical diversity in narrative production by normal healthy adults was investigated and no age-related disadvantage for TTR was noted across the different age groups of participants. Findings of the current study further resembled that of Kemper and Sumner’s study, where, in fact, an age-related advantage of TTR was found (i.e., older adults tended to produce narratives with a higher TTR compared to younger adults), although this advantage did not reach statistical significance in the current study. Such an age-related advantage of TTR for the older participants may be at least partly attributed to their greater lexical knowledge. The North American Adult Reading Test — short form (NAART35) was administered to each participant as a measure of lexical knowledge, and a significant difference was found between the raw score of the two age groups. Specifically, the older age group had a higher mean raw score than the younger age group. Therefore, the broader lexical knowledge of the older participants may have allowed them to use a greater variety of vocabulary during narrative production, resulting in a higher TTR than the younger age group.
4.1.1.2 Age and verb-noun ratio.

The verb-noun ratios of the narratives between the two age groups were also not significantly different. In fact, the mean verb-noun ratios of the two age groups were very similar. Such a finding was consistent with the results of other studies, where verb-noun ratio in narrative production was investigated and no age difference was found (Marini et al., 2005).

4.1.1.3 Age and coherence.

Based on Joanette and Goulet’s (1990) protocol, coherence was measured by the total number of irrelevance, non-progression, and contradiction errors in a given piece of discourse. Although the older participants on average produced more coherence errors in their narratives than the younger participants, the difference between the two groups was not significant. This result was consistent with that of previous research, where picture sequence description tasks were employed in order to investigate age-related difference in the degree of coherence in narrative production by normal healthy adults. For example, in Marini et al.’s (2005) study, coherence errors such as off-topic and tangential speech were measured. However, even the participants from the oldest age group (75 to 84) did not perform differently than those from the younger age groups. On the other hand, it is worth noting that in both the present and Marini et al.’s studies, a floor effect was noticed for the coherence error measure. That is, there were in general very few coherence errors in the narratives of both age groups. In light of the observation, interpretation of the results should be cautious, since the floor effect may compromise the validity and reliability of the statistical analyses of the raw data.
4.1.2 The effects of age on cohesion and content richness

The second hypothesis of the current study predicted that age would have a negative effect on, first, cohesion and, second, content richness in narrative production, albeit for different reasons. However, findings of the current study did not support this hypothesis.

4.1.2.1 Age and cohesion.

Cohesion was measured by the accuracy and clarity of the use of reference. The total number of incomplete, inaccurate, or ambiguous references was coded and recorded as a measure of cohesion error. Despite the relatively large amount of evidence that supports an age-related decline in cohesion in narrative production as measured by reference errors (Juncos-Rabadán et al., 2005; Marini et al., 2005; Pratt et al., 1989; Ulatowska et al., 1986), the mean numbers of cohesion errors of the younger and the older age group of the current study were found not to be significantly different. The finding of the current study also provided evidence against the assumption that referencing is very closely associated with cognitive abilities such as working memory (Ulatowska et al., 1986; Pratt et al., 1989). In Chapter 1, working memory was speculated to be important to the maintenance of cohesion through referencing. Specifically, it has been reported that individuals with a smaller working memory span size tended to produce more cohesion errors in narrative production (Pratt et al., 1989). Results of the current study, however, showed that although the older age group had a significantly lower sentence span and counting span size than the younger age group, the number of cohesion errors did not differ significantly. This suggested that working memory is not solely responsible for supporting reference use; some other mental processes
must also be at least partially contributing to the production of complete and unambiguous references.

It is important to note that although the difference between the mean total number of cohesion errors of the two age groups was not significantly different, the narratives of the participants in the older age group did on average contain more cohesion errors than those of the participants in the younger age group. The lack of statistically significant difference could be due to some inherent limitations of the study such as small sample size. Similar to coherence errors, a floor effect was also noted for the cohesion measure. Participants of both age groups in general produced very few cohesion errors. Therefore, the failure to detect any significant main effect of age on cohesion could be due to a floor effect.

4.1.2.2 Age and content richness.

Content richness of each narrative was measured in terms of proposition density, which was the total number of propositions in a narrative divided by the total number of words. Although there is abundant evidence in the current literature that supported a negative effect of age on content richness in narrative production (Duong and Ska, 2001; Kemper et al., 2001; Kemper and Sumner, 2001; Marini et al., 2005; Pratt et al., 1989; Wright et al., 2005), the current study did not find a significant difference between the proposition density of the younger and the older age groups. In fact, the older participants were found to have a slightly higher proposition density than the younger participants, albeit insignificantly so. Also, many of the aforementioned studies that found an age-related difference in content richness in narrative production employed a core proposition analysis where the number of main events of a story produced in a narrative, rather than the proposition density of a
narrative, was measured (Duong and Ska, 2001; Marini et al., 2005; Pratt et al., 1989; Wright et al., 2005). While core proposition analysis and proposition density are both methods used to determine content richness of a piece of narrative, they are conceptually and methodically different measures. Therefore, just because age was found to have an effect on the number of core propositions produced does not necessarily mean that age would also have an effect on proposition density.

4.1.3 The effects of task on cohesion

It was also hypothesized that participants, especially older adults, would commit more cohesion errors in the video condition than in the picture sequence condition. This was speculated to be due to the higher demand the video condition placed on the working memory system, which could compromise the use of clear references, given the close relationship between working memory and referencing as discussed previously. However, results of the current study did not support this hypothesis. Not only was the difference between the mean numbers of cohesion errors of the two task conditions insignificant, narratives from the picture sequence condition actually contained slightly more cohesion errors than those from the video condition. One possible explanation for such an observation is that there was a shared physical context between the examiner and the participant in the picture sequence condition but not in the video condition. Recall that a difference between the two task conditions was that participants were allowed to look at the stimulus while they were producing the narratives in the picture sequence condition but not in the video condition. Because the picture sequence was in front of both the participants and the examiner during the production of the narrative, the participants may have explicitly or implicitly assumed that the examiner shared their knowledge of which character, object or event in the pictures
they were talking about. With this assumption, participants may have become less vigilant or
invested less effort in maintaining the accuracy and clarity of the referencing devices they
used, thus leading to more cohesion errors in the narratives elicited from the picture sequence
condition. It is also important to remember that there was a floor effect in general in the
performance of all participants in terms of committing cohesion errors. This may have
interfered with the statistical analysis of the data, masking any significant difference between
the cohesion of the narratives elicited from the two task conditions.

4.1.4 The effects of task on coherence and content richness

Finally, it was also hypothesized that participants would commit more coherence
errors and produce narratives that were less rich in content in the picture sequence condition.
This was speculated to be due to the potentially higher cognitive demands imposed on the
participants in the picture sequence condition compared to the video condition in terms of
inference generation and symbol processing. The greater demand of inference generation and
symbol processing was anticipated to lead to a greater likelihood of misunderstanding and
misconception of the stimulus, which ultimately adversely affect both the degree of
coherence and the content richness of narratives elicited from the picture sequence stimulus.

The findings did not support this hypothesis: no significant difference was found
between narratives elicited from the two task conditions in terms of both the total number of
coherence errors and proposition density. In fact, in contrast to the original prediction,
slightly more coherence errors were actually noted in the video condition than in the picture
sequence condition. Also proposition density was found to be slightly higher in the picture
sequence condition. One possible explanation for the lack of significant difference between
task groups is that only adults without language or cognitive impairments were recruited to participate in this study. As a result, participants were able to perform well within normal limits in the generation of inferences and processing of symbols despite the possibly greater cognitive demands imposed by the picture condition. This may have led to a ceiling effect in terms of inference generation and symbol processing. As a result, the narratives of both task conditions contained very few coherence errors in general and had proposition densities that were insignificantly different.

Another possible explanation of the lack of significant difference in proposition density between the two task conditions involves the way that proposition density was calculated. Proposition density was determined by dividing the total number of propositions by the total number of words (TNW) in a narrative. Mathematically speaking, the higher the TNW, the smaller the proposition density value is likely to be. In section 3.2.1, it was shown that narratives of the video condition had a significantly higher TNW than those of the picture sequence condition. Therefore, this may at least partially account for the lower mean proposition density of the video condition, although insignificantly so.

4.2 Other Research Findings: Insights into Dependent Variable Measures

Apart from the findings discussed above, there were also results from this study that were not originally anticipated but nonetheless provided important insights into different aspects of the current study, such as the protocol used to measure the dependent variables. The following sections discuss three of these findings, namely, the relationship among task, total number of words and type-token ratio, the floor effects noted for the cohesion and
coherence measures, and the results of the core proposition analysis of the Cowboy's Story narratives.

4.2.1 Type-token ratio: the effect of total number of words

Type-token ratio (TTR) was found to be higher for narratives elicited from the picture sequence condition as opposed to the video condition. However, the result of a post-hoc analysis showed that the total number of words (TNW) of the narratives was significantly higher for narratives in the video condition than those in the picture sequence condition and that TNW was a highly significant covariate of the effect of task on TTR. Specifically, task no longer had a significant effect on TTR once TNW was included as a covariate in the analysis.

The sensitivity of TTR to TNW raised the question of whether TTR was the ideal measure for lexical diversity. Evidence in the literature of child language also suggested that TTR may not be sensitive enough to differentiate individuals with different levels of lexical diversity in spontaneous language samples (Watkins, Kelly, Harbers, & Hollis, 1995). In light of TTR's sensitivity to sample size, researchers have explored other measures that estimate lexical diversity, such as number of different words (NDW) and D value.

4.2.1.1 Number of different words.

NDW has become one of the most preferred measures of lexical diversity in child language research. According to Wright and colleagues (2003), measuring and comparing NDW across language samples is essentially comparing TTR of these samples with TNW controlled for, because in the latter analysis, NDWs of the narratives are simply divided by a common denominator. Therefore by measuring NDW, researchers may avoid the
confounding influence of TNW on TTR, thus obtaining a more valid and reliable measurement of lexical diversity.

4.2.1.2 D value.

Originally developed by McKee, Malvern and Richards (2000), $D$ is a measure of lexical diversity derived from TTR by accounting for its sensitivity to sample size variation. This is achieved by taking into consideration the probability of new vocabulary being introduced into a language sample that progressively gets longer. If a language sample is transcribed in the CHAT format of the CHILDES project, the $D$ value of a sample can be computed using the `vocd` utility of the CLAN language sample analysis program (MacWhinney, 2000).

Studies have been conducted to examine the validity of $D$ as a measure of lexical diversity in both the area of child language and acquired language disorders. Owen and Leonard (2002) used $D$ to compare the lexical diversity of children with SLI to that of their same-age and younger normal-developing peers. $D$ appeared to be sensitive to different developmental levels, as younger children were found to have a smaller $D$ than older normal developing children and children with SLI. They also found that the effect size of sample size was significantly larger for TTR than for $D$. Wright et al. (2003) investigated the ability of TTR and $D$ to differentiate fluent and nonfluent aphasia. Nine individuals with fluent aphasia and nine with nonfluent aphasia were asked to produce conversation and narrative samples. Both TTR and $D$ were calculated for each narrative. Results of the study showed that the experimental groups differed significantly by $D$ but not by TTR. In summary, there is some evidence that supports the use of $D$ as a measure of lexical diversity over TTR. More
research is needed to further investigate the validity and reliability of $D$ when used in discourse research on the adult population.

4.2.2 Total number of words as a primary dependent variable

Because a main effect of task was found on TNW, it proved to be a primary measure of interest itself. The higher TNW in the video condition may be attributed to the higher level of transparency of the stimulus. Recall that the picture sequence condition was speculated to impose a greater demand on the participants in terms of symbol processing. This is because the picture sequence was arguably a more opaque representation of a story, where the gist of the story was not as readily available to the participants as it was in the video condition. Therefore, participants possibly needed to make more effort in processing the pictorial materials to make sense of the story before they could produce the narrative about it. While it was discussed in section 4.1.4 that the greater demand on symbol processing in the picture sequence condition did not negatively affect coherence and content richness of the resulting narratives, the more transparent nature of the video stimulus might have stimulated the participants to produce narratives that were more elaborate and contained more details, thus increasing the length of the narratives.

4.2.3 Floor effect of cohesion and coherence error

Because only normal healthy adults were included in the current study, there were floor effects noted on several dependent measures. The existence of floor effects in the cohesion and coherence measures may indicate that the scheme used in the current study is not very sensitive in identifying cohesion and coherence errors, and thus, may not be an ideal tool to differentiate the level of cohesion and coherence of the narratives between the age and
task groups. Indeed the scheme used in this study was originally employed by Joanette and Goulet (1990) in their study that explored the characteristics of narrative discourse of individuals with RHD. Therefore, while appropriate for clinical populations, this scheme may actually be unsuitable for studies that involve only non-clinical populations. There are other cohesion and coherence coding schemes that have been used by researchers and clinicians and may potentially be more successful in differentiating the different levels of cohesion and coherence of narratives produced by different experimental groups. One of these potential schemes is the Narrative Scoring Scheme (Miller & Iglesias, 2008).

Originally developed to investigate reading and academic achievement of students learning English as a second language, the Narrative Scoring Scheme is part of the Systematic Analysis of Language Transcript software (Miller & Heilmann, 2009; Miller & Iglesias, 2008). Since its development, it has been used by clinicians in assessing the oral language skills of preschoolers and children in their early school-age years (SALT software, n.d.). The NSS is essentially a holistic measure of the quality and completeness of narratives that incorporates the analysis of cohesion, coherence, story grammar and content richness. Each narrative is given a subjective rating of zero to five in seven categories, namely, introduction, character development, mental states, referencing, conflict/resolution, cohesion and conclusion. A score of zero represents very poor performance while a score of five represents proficiency in that particular category. The score of each category can then be summed, resulting in a total score for the narrative. Although no use of the NSS on the adult clinical or non-clinical populations has been documented so far, and thus, the reliability of this subjective 5-point rating scale system has yet to be established with the adult population, it is a potential tool for the assessment of discourse competence of adults.
4.2.4 Core proposition analysis of the Cowboy’s Story narratives

In order to compare the narratives of the participants of the current study with those of the normal healthy control participants in Joanette and Goulet’s study, a core proposition analysis was conducted on the Cowboy’s Story narrative of each participant, be it elicited in the picture sequence or video condition. This was accomplished by evaluating the propositions of each Cowboy Story with respect to the list of core propositions identified for the Cowboy’s Story in Joanette and Goulet’s (1990) study such that the percentage of core propositions produced for each Cowboy’s Story narrative was computed. The core proposition analysis was only carried out for the Cowboy’s Story narratives because a list of core propositions has yet to be developed for the Fisherman’s Tale.

While the core propositions for the Cowboy’s Story narratives produced by the participants in the current study were similar to those produced by the normal healthy control group in Joanette and Goulet’s (1990) original study (although note that participants in this study produced substantially more core propositions), one unexpected finding was a main effect of task on the percentage of core propositions produced. Specifically, the Cowboy’s Story narratives elicited from the video condition on average contained more core propositions than those from the picture sequence condition. However, recall that a main effect of task was not found on proposition density for the Cowboy’s Story narratives. The discrepancy between these two findings further draws attention to the fact that proposition density and core proposition analysis are different measures of content richness. Specifically, proposition density measures content richness of a piece of discourse in terms of how much information is conveyed and how efficiently such information is conveyed. The relevance of the information to the topic of the discourse task at hand is not taken into consideration
during the computation of proposition density. Core proposition analysis, on the other hand, measures content richness in terms of the amount of relevant information conveyed. In the case of narrative production, the more information conveyed that is relevant (i.e., the higher the number of core propositions produced), the more complete the narrative is considered to be. However, while relevance and completeness of the information are taken into account in a core proposition analysis, the efficiency of how such information is conveyed is not considered. Therefore, given the nature of core proposition and proposition density analysis, these two measures are actually qualitatively different but complementary evaluations of content richness.

The main effect of task on the percentage of core propositions produced also had crucial implications on the evaluation of the research hypotheses discussed earlier. Recall that part of the fourth research hypothesis anticipated content richness to be significantly higher on average in the narratives of the video condition than those in the picture sequence condition (see section 1.6 and 4.1.4). If content richness were measured by the percentage of core propositions produced instead of proposition density, this part of the fourth research hypothesis might have been fulfilled by the finding that narratives in the Cowboy's Story video condition contained more core propositions. However, it is important to remember that the main effect of task on the percentage of core propositions produced was only observed for the Cowboy's Story narratives, since a core proposition analysis was not conducted for the Fisherman's Tale narratives. Therefore a firm conclusion cannot be drawn until a core proposition analysis is also conducted on the Fisherman's Tale narratives.

Some researchers have investigated ways to incorporate a core proposition analysis with other measures of narratives such as macrostructure analyses in order to have a more
comprehensive evaluation of the content richness and structure of narratives. For example, in a pilot study that examined story-retell ability of the veterans from the Vietnamese war who sustained a penetrating head injury (PHI), Coehlo and colleagues (Coehlo, Le, Mozeiko, Kruger, & Grafman, 2009) developed a measure to investigate what they termed “story goodness” in narrative production. Story goodness is a measure that combines a story grammar analysis with a story completeness analysis. For the story grammar measure, the macrostructure of each episode produced for the story is examined and the number of t-units within each episode is measured. For the story completeness analysis, the number of pre-identified core propositions produced is measured. The results of these two analyses for each narrative are then plotted onto a graph, where the x-axis is the story completeness score and the y-axis the story grammar score. Story goodness of a narrative is then rated by determining which quadrant captures the story goodness point. The story goodness point of each participant belonging to the same experimental group can be plotted onto the same graph, and the graphs of the different groups can be compared visually.

4.3 Limitations

Firstly, in this study, participants were asked to produce two narratives, one from the Cowboy’s Story stimulus and one from the Fisherman’s Tale stimulus. Despite the effort to have the two stories matched, evidence that suggested the two stories were not equivalent was found. This finding must be taken into consideration in the interpretation of the data, since the potential differences between the two stories may have masked some significant effects of age and/or task condition on the dependent measures. It is important to note that analysing the two stories separately was also problematic since the sample size decreased by
half for each analysis. This decreased the general statistical power of the analysis and may have masked potentially significant effects.

Secondly, despite the findings in the literature that support an age difference in at least certain aspects of narrative production, no significant effect of age was found on any of the dependent measures. The discrepancy between the findings of the current study and those of previous research could be due to a difference in the age range of the older participants. The mean age of the older participants of the current study was 69.6, with a range of 63 to 81. However, in studies that found age-related differences in narrative production, the changes did not become apparent until an even older age. For example, in Ulatowska et al.’s (1986) study on referencing as a cohesive device in spontaneous discourse, a decline in referencing only started to emerge in their second oldest age group, which had an age range of 64 to 76. The decline only became apparent in their oldest age group, which had an age range of 77 to 92. In Marini et al.’s (2005) study, content richness of the narratives produced by their participants was compared across the five continuous age groups ranging from 20 to 84. It was found that the content richness of only the oldest age group (75 to 84) and not even the second oldest age group (60 to 74) differed significantly from that of the three younger age groups. Duong and Ska’s study (2001) on age-related changes in content richness provided further evidence that young-old adults performed differently than old-old adults. Specifically, they found that young-old adults (65 to 74) produced narratives that were richer in content than old-old adults (75 to 84). In summary, comparing the participants of the current study to those in previous research, one may argue that the older participants in the current study were too young to demonstrate significant age-related changes in narrative production.
Another variable that potentially has an impact on narrative production but was not considered and controlled for in this study was the years of formal education. Performance on narrative production may be argued to be closely tied to the amount of formal education one has received. For example, the higher one's education level is, the more likely one would use a greater variety of vocabulary, produce narratives that are more cohesive and coherent and richer in content, given their presumably broader experience with text processing. For example, the difference between the percentage of core propositions produced by Joanette and Goulet's participants vs. those in the current study could be at least partially due to the lower overall education in the former group (approximately 8 vs. 18 years). Given the speculated close relationship between education level and narrative production, it is important to take into consideration when interpreting the results that education was not controlled for in this study. When generalizing the findings of the current study, there is also the question of whether the participants of the current study were representative of the general population of normal healthy adults in terms of education level.

4.4 Clinical Implications

The results of the current study add to the pool of normative data in the current literature against which clinicians can compare their own clients. In addition, two findings of this study may have potentially important clinical implications. Firstly, while task did not have any main effects on any dependent measures, results of this study suggested that different stories may influence the resulting narrative differently in terms of measures such as lexical diversity and cohesion. Therefore when clinicians are comparing the performance in narrative production across clients or comparing the narratives produced by the same client at different times in the recovery process, it appears to be more desirable to use the same story
stimulus. Also if clinicians and researchers intend to develop some kind of norms for narrative production of different adult clinical and non-clinical populations, it appears to be more appropriate for these norms to be story-specific. In addition, these findings lend support to recommendations such as those of Brookshire and Nicholas (1994), who suggested that discourse measures should be taken across multiple stimuli (e.g., different tasks and stories). Secondly, a video retell task may potentially elicit significantly longer and more complete narratives in terms of the total number of words (TNW) and core propositions than a picture sequence retell task. However, the effect of task on these as well as other aspects of narrative needs to be further explored with clinical populations.

4.5 Directions for Future Research

Because only adults without language or cognitive impairments participated in this study, the difference in the difficulty level of the two tasks conditions may not have been great enough to cause a difference between the performances on the two tasks. As a result, floor effects were noted in the dependent measures of the narratives in both task conditions. As discussed in section 4.2, there are other possible methods to measure the dependent variables that may be more appropriate for this study or may yield other valuable information about the effects of age and task on narrative production. Therefore, the data gathered in this study can be reanalyzed with some of these methods. For example, lexical diversity can be measured by the number of different words or the $D$ value instead of type-token ratio since the former two measures are less sensitive to the length of the narratives in terms of total number of words. Cohesion, coherence and content richness can be measured using the Narrative Scoring Scheme (Miller & Iglesias, 2008). Instead of proposition density, a list of core propositions can be developed for the Fisherman’s Tale as well such that a core
proposition analysis can be used to measure content richness. Lastly, content richness may also be examined by adopting the story goodness protocol developed by Coehlo et al. (2009).

Another important step would be to study these aspects of narrative production in clinical populations such as individuals with aphasia, right-hemisphere damage (RHD) and traumatic brain injury (TBI) and to compare their performance to the non-clinical population. Because the hallmark symptom of these aforementioned clinical populations is cognitive-linguistic deficits, and given the relationship between cognitive functioning and some of the dependent measures of this study discussed in Chapter 1, the performance of individuals of these populations on the dependent measures of this study is of great interest.

4.6 Conclusion

Narrative production is an inseparable part of our daily lives. It also serves as an invaluable tool in the assessment and treatment of developmental and acquired language disorders. Because acquired language disorders secondary to brain injury or degenerative disorders are especially common among older adults, it is important to distinguish the effects of normal aging from the effects of these acquired conditions on narrative production. Also, because narratives can be elicited by using different tasks and visual materials, it is also important to explore how the variation in tasks and materials used affect the resulting narratives. This study was designed to investigate the effects of age and the use of different visual stimuli (picture sequence vs. video) on narrative production by normal healthy adults.

Results of the current study showed that neither age nor tasks had significant effects on lexical diversity, verb-noun ratio, cohesion, coherence and content richness as measured by proposition density in narrative production. However, a video stimulus tended to elicit
longer and more complete narratives, at least in the case of one story. The choice of story also appeared to influence the resulting narratives. Although several limitations of the current study were noted, these results nevertheless have some important clinical implications and provide directions for future clinical research. All in all, more research on narrative production is needed in order to further explore the different discourse characteristics of both clinical and non-clinical adult populations.
5. References


APPENDIX A: Ethics certificate

The University of British Columbia
Office of Research Services
Behavioural Research Ethics Board
Suite 102, 6190 Agronomy Road,
Vancouver, B.C. V6T 1Z3

CERTIFICATE OF APPROVAL - MINIMAL RISK

<table>
<thead>
<tr>
<th>PRINCIPAL INVESTIGATOR:</th>
<th>INSTITUTION / DEPARTMENT:</th>
<th>UBC BREB NUMBER:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barbara A. Purves</td>
<td>UBC/Medicine, Faculty of/Audiology &amp; Speech Sciences</td>
<td>H08-03011</td>
</tr>
</tbody>
</table>

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:

<table>
<thead>
<tr>
<th>Institution</th>
<th>Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>UBC</td>
<td>Vancouver (excludes UBC Hospital)</td>
</tr>
</tbody>
</table>

Other locations where the research will be conducted:
For older adults, experimental sessions may be conducted at suitable locations of their choice (e.g., their community centre, senior organizations)

CO-INVESTIGATOR(S):
Clinton Yin Hang Tsang

SPONSORING AGENCIES:
N/A

PROJECT TITLE:
Effects of age and visual stimuli on narrative production
CERTIFICATE EXPIRY DATE: March 9, 2010

<table>
<thead>
<tr>
<th>DOCUMENTS INCLUDED IN THIS APPROVAL:</th>
<th>DATE APPROVED:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protocol:</td>
<td>March 9, 2009</td>
</tr>
<tr>
<td>Picture/Video Stimuli</td>
<td>N/A March 6, 2009</td>
</tr>
<tr>
<td>Working Memory Span Tasks</td>
<td>N/A February 11, 2009</td>
</tr>
<tr>
<td>NAART-35</td>
<td>N/A March 6, 2009</td>
</tr>
<tr>
<td>Consent Forms:</td>
<td></td>
</tr>
<tr>
<td>Subject Information and Consent Form</td>
<td>N/A March 6, 2009</td>
</tr>
<tr>
<td>Advertisements:</td>
<td></td>
</tr>
<tr>
<td>Recruitment Poster</td>
<td>N/A March 6, 2009</td>
</tr>
<tr>
<td>Questionnaire, Questionnaire Cover Letter, Tests:</td>
<td></td>
</tr>
<tr>
<td>Participant questionnaire</td>
<td>N/A March 6, 2009</td>
</tr>
</tbody>
</table>

The application for ethical review and the document(s) listed above have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.

Approval is issued on behalf of the Behavioural Research Ethics Board and signed electronically by one of the following:

Dr. M. Judith Lynam, Chair  
Dr. Ken Craig, Chair  
Dr. Jim Rupert, Associate Chair  
Dr. Laurie Ford, Associate Chair  
Dr. Anita Ho, Associate Chair
APPENDIX B: Participant information and informed consent

The University of British Columbia
Faculty of Medicine
School of Audiology and Speech Sciences

PARTICIPANT INFORMATION AND CONSENT FORM

Effects of Age and Visual Stimuli on Narrative Production

Principal Investigator: Dr. Barbara A. Purves, PhD, Assistant Professor
School of Audiology and Speech Sciences
UBC Faculty of Medicine
(604) 822-2288

Co-Investigator: Clinton Yin Hang Tsang, M.Sc. Candidate
School of Audiology and Speech Sciences
UBC Faculty of Medicine

Background:
Narratives are a key feature of everyday life. The study of narratives is an invaluable tool to comprehensively investigate how individuals use language (e.g. use of vocabulary, sentence structure and discourse). Research in narratives in cognitively intact adults may also yield results that can serve as baseline for comparison with the clinical population such as people who suffer from acquired brain injuries. However, despite these tremendous values of studying narrative production, research in this area remains scarce. Among the existing research studies, two factors, age and the type of stimuli used to elicit narratives, have been shown to influence narrative production. This study further investigates the effect of aging and the use of picture sequence vs. video as visual stimuli on verb use, cohesion and coherence, and content richness in narrative production by cognitively intact adults.

Why Are You Invited:
You are invited to participate in this study because you are an English-speaking adult with no current and previous cognitive and language impairment. This research is being carried out as part of Clinton Yin Hang Tsang’s Master’s degree.
Study Description:

Your participation in this study will involve: a) completing a short questionnaire, b) completing a reading test; c) completing a series of working memory tasks; and d) telling two stories, one based on an eight-frame picture sequence shown on paper and the other based on a short video shown on a computer screen. The researcher, Clinton Yin Hang Tsang, will meet with you at the Acquired Language Disorders Lab located in the School of Audiology and Speech Sciences at the University of British Columbia, for a one-time session of approximately one hour.

Possible Benefits and Risks of Participation:

There are no known risks as a result of participating in this project. If you become fatigued during the tasks in this study, you will be allowed to take a break or discontinue the testing session. You will not be asked to complete any activities against your will. If you consent to participating in this study, you may benefit from learning about the research process (e.g. what it is like to be a study participant), and, upon completion of the study, finding out more about the topic of narrative production. We will also send you a summary of the study’s overall findings upon request. Please note that your individual scores or data will not be disclosed.

Honourarium and Reimbursement:

Participants will receive an honourarium of ten dollars for participating in the study. In addition, a reimbursement of up to ten dollars will be offered to subjects for expenses incurred in attending the experimental session.

Confidentiality:

All information pertaining to your participation will be kept strictly confidential. The data and records will be stored in a locked cabinet or password-secured computer in the Principal Investigator’s research lab. Only Mr. Tsang and members of his thesis committee will have access to information that could identify you.

Data may be used for educational purposes or at scientific meetings. All identifying information will be removed prior to such use.

Participation:

Participation in this study is entirely voluntary. If you choose to participate in this study, you will be asked to sign this consent form. Please remember that even if you do consent to participation, you may still withdraw from the study at any time.

If you have any questions or would like further information about this project, please feel free to contact the Co-Investigator Clinton Yin Hang Tsang or the Principal Investigator Dr. Barbara A. Purves at (604) 827-3042 or (604) 822-2288.

If you have any concerns about your rights as a research subject and/or your experiences while participating in this study, you may contact the Research Subject Information Line in the University of British Columbia Office of Research Services at (604) 822-8598.
Consent to Participate

Your signature below indicates that you have received a copy of this consent form for your own records.

Your signature indicates that you consent to your participation in this study.

I have read and understood the subject information and consent form of this research study. I, __________________________ agree to participate in this study.

Your signature: __________________________________________

Witness’ name: ___________________________________________

Witness’ signature: _______________________________________

Date: ___________________________________________________

Thank you very much!
Debris
Simile
Subtle
Bouquet
Colonel
Rarefy
Gist
Corps
Hors D’oeuvre
Sieve
Hiatus
Gauche
Zealot
Paradigm
Façade
Cellist
Indict
Détente
Impugn
Aeon
Epitome
Reify
Indices
Assignate
Topiary
Caveat
Leviathan
Quadruped
Sidereal
Abstemious
Beatify
Goaled
Demesne
Syncope
Ennui

APPENDIX C: NAART35 word list
APPENDIX D: Fisherman's Tale
APPENDIX E: Sample transcript

@Begin
@Languages: en
@Participants: PAR Participant, EXP Experimenter
@ID: E104
@Birth of PAR: 03-SEP-1941
@Coder: Clinton_Tsang
@Activity: Cowboy_Picture
*PAR: one day a cowboy # who had been out on the range all day checking the fences.
*PAR: and he was pretty tuckered.
*PAR: been a long long hard # day in the hot sun.
%err: $REF:GAP
*PAR: and he returned home.
*PAR: he was so tired.
*PAR: and he sat down to rest.
*PAR: and even his horse was exhausted.
*PAR: pretty soon he was asleep.
*PAR: then a little boy came along with his toy horse on a leash.
*PAR: and he saw the man sleeping there the cowboy and his horse.
*PAR: and the cowboy had his horse's reins in his hand.
*PAR: and the boy thought +"./.
*PAR: +" ah I'm going to play a little trick here.
*PAR: so he took some shears.
*PAR: and he cut the reins # that the man was holding so the horse was free to go.
*PAR: then he proceeded to take # the leash that was attached to his toy horse and put it in
the man's hand.
*PAR: and then he left along with the man's horse.
*PAR: after a while the man woke up.
*PAR: and when he saw in his hand a miniature horse he was totally amazed and could not
believe what had happened.
*PAR: the end.
@End

Key to sample transcript:
*PAR: participant's utterance  
%err: error tier
#: pause
+"/: direct speech follows
$REF:GAP: reference error (see Appendix G)
APPENDIX F: Definitions and CLAN codes for cohesion and coherence errors

Cohesion error:

$REF:GAP — reference errors
- Ambiguous referents when using pronouns
- Lack of antecedents when using pronouns
- Missing subjects in a main clause
- Ambiguous use of determiners (e.g. “a” and “the”)

Coherence error:

$COR:CONT — contradiction errors
- Occurs when new information contradicts with previously given information.

$COR:NPRO — non-progression errors
- Repetition of previously given information without the addition of new information.

$COR:IRR — irrelevance errors
- New information is given without any relation either explicitly or implicitly to previously given information or the main theme of the story, which may or may not result from a misconception or misunderstanding of the visual stimuli.
APPENDIX G: Sample proposition analysis of transcript in Appendix E

1. Quality of(cowboy, out)
2. On(P1, range)
3. Extent of(day, all)
4. Extent of(P2, P3)
5. Check(A: cowboy, O: fences)
6. Qualify(P4, P5)
7. Quality of(cowboy, tuckered)
8. Qualify(P7, pretty)
9. Quality of(day, hard)
10. Extent of(P9, long)
11. Quality of(sun, hot)
12. In(P10, P11)
13. Is(it, P12)
14. Returned(A: cowboy, O: home)
15. Quality of(cowboy, tired)
16. Qualify(P15, so)
17. Sit down(A: cowboy)
18. Rest(A: cowboy)
19. To(P17, P18)
20. Possess(A: cowboy, O: horse)
21. Quality of(P20, exhausted)
22. Qualify(P21, even)
23. Qualify(soon, pretty)
24. Quality of(cowboy, asleep)
25. Qualify(P24, P23)
26. Quality of(boy, little)
27. Qualify(horse, toy)
29. Come(A: P26)
30. Qualify(P29, along)
31. With(P30, P28)
32. On(P31, leash)
33. Qualify(P32, then)
34. And(cowboy, P20)
35. Sleep(A: P34)
36. Qualify(P34, there)
37. See(A: P26, O: P36)
38. Possess(A: cowboy, O: hand)
39. Part of(P20, reins)
40. Possess(A: cowboy, P39)
41. In(P40, P38)
42. Quality of(trick, little)
43. Play(A: I, O: P42)
44. Qualify(P43, here)
Think(A: boy, O: P44)

Number of(shears, some)

Take(A: boy, O: P46)

Hold(A: man, O: reins)

Cut(A: boy, O: P48)

Quality of(horse, free)

Go(A: horse)

To(P50, P51)

So(P49, P52)

And(P47, P53)

Attach(A: $, O: leash, G: P28)

Take(A: boy, O: P55)

Put(A: boy, O: P55, G: P38)

And(P56, P57)

Proceed(A: boy)

To(P59, P58)

Qualify(P60, then)

Leave(A: boy)

Possess(A: man, O: horse)

With(P62, P63)

Qualify(P64, along)

Qualify(P65, then)

Wake up(A: man)

After(P67, while)

Possess(A: man, O: hand)

Quality of(horse, miniature)

See(A: man, O: P70)

In(P71, P69)

Quality of(man, amazed)

Qualify(P73, total)

Happen(O: what)

Believe(A: man, O: P75)

Negate(P76)

And(P74, P77)

When(P78, P72)
APPENDIX H: Core propositions of the Cowboy's Story from Joanette and Goulet (1990)

1. Arrive, Cowboy
2. On, Cowboy, Horse
3. In, P1, village
4. Get off, Cowboy, Horse 1
5. Of, Horse 1, Cowboy
6. Tie, Cowboy, Horse 1
7. Tired, Cowboy
8. Sit, Cowboy, Bench
9. Sleep, Cowboy
10. Hold, Cowboy, Horse 1
11. Arrive, Boy
12. Small, Boy
13. Hold, Boy, Horse 2
14. Small, Horse 2
15. Of, Horse 2, Wood
16. Play, Trick, Boy
17. Cut, Boy, Bridle
18. Of, Bridle, Horse 1
19. With, P17, Scissors
20. Sleep, Cowboy
21. Tie, Boy, Horse 2
22. Of, Horse 2, Boy
23. To, P21, Hand
24. Of, Hand, Cowboy
25. Instead of, P21, Horse 1
26. Run away, Horse 1
27. Run away, Boy
28. And, P26, P27
29. Wake up, Cowboy
30. Notice, Cowboy, P31
31. Have, Cowboy, Horse 2
32. Overturn, Cowboy, Bench