WHO'S WHO: MEMORY UPDATING AND CHARACTER REFERENCE IN CHILDREN'S NARRATIVES

by

Cristy L. McNiven

B.A., The University of British Columbia, 2007

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

MASTER OF SCIENCE

in

The Faculty of Graduate Studies

(Audiology and Speech Sciences)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

March 2010

ABSTRACT

This study investigated whether individual differences in working memory updating and adequate reference to story characters in narrative discourse were related in a group of typically-developing children from kindergarten through second grade. It also documented developmental trends in both of these abilities, and examined factors that may have contributed to the difficulty of clearly referring to story characters.

The results indicate that the ability to update working memory is related to referential adequacy in children 5 to 8 years of age, with all three updating tasks (visual, auditory, and verbal) correlating moderately and significantly with referential adequacy scores. The participants were most successful when maintaining reference to story characters, and had more difficulty when introducing and reintroducing characters. An analysis of the three referential functions showed that updating was significantly related to children's adequate maintenance and reintroduction scores. The strongest relationship occurred between updating and maintenance, which was unexpected, and may be specific to this developmental window when children are continuing to develop the linguistic skills required for clearly referring to characters. The bivariate correlations among the updating tasks were all moderate and significant. The consistent correlations amongst the updating tasks, as well as the significant moderate correlation between the visual updating task and referential adequacy, suggest that domain-general resources are involved in updating. This study is the first to investigate the possible link between updating abilities and language production in children. Although these findings are preliminary, they point to a relationship between updating and adequate reference to story characters in narrative production in 5- to 8-year-old children.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I owe much thanks to Paola Colozzo and Judith Johnston for their supervision on this project. Judith, thank you for all your time and great ideas as we designed the study, and for the hours you have committed to helping me in your retirement. Paola, I am most grateful for the many, many hours you spent working with me during the past year. To my committee member, Carolyn Johnston, thank you for bringing your expertise to the table.

To the children, teachers, and principals from the Maple Ridge School District who participated in my project, it was not possible without you. Thank you to Carol Woodworth who was most helpful in getting the data collection started. A big thank you to Heidi Ostberg, Sarah Wright, and Heather Morris, who were a huge help with data collection and analysis. Finally, to my friends and family who leant their support along the way, I owe you big time.

1. INTRODUCTION

Creating and maintaining clear reference to story characters is an important part of successful narration and a complex task for young children to master. A narrative text (either oral or written) is a sequence of sentences that relates events and the temporal-causal relationships between them (Halliday & Hasan, 1976). Karmiloff-Smith (1980) emphasized the dynamic, on-going, nature of a narrative production: it is a text in use that requires constant updating and modification of a discourse representation in order to result in a successful communicative event. Within the narrative, reference to people, places, and objects can be accomplished by providing a proper name, using a noun phrase, or utilizing a cohesive element that links to another more specific mention in the text. The speaker must be able to select the correct form and manage references to the story characters as they enter and re-enter the action in the story. From the listener's perspective, the speaker must clearly express whom he is talking about, otherwise identities can become confused, which may compromise a story's coherence. In order to clearly refer to story characters, the young language learner must acquire the various referring expressions, understand the distinctions between them, and learn the rules governing their use. The present study focused specifically on the ability of children to clearly refer to story characters, and on the updating of working memory that is required to adequately maintain character identities during a narrative production task.

1.1. Referring Expressions

As the events in a narrative unfold, the speaker must ensure that the identity of the various story characters is clear for the listener. The forms necessary to succeed in this task differ depending on whether the storyteller is introducing a new character, maintaining reference to a current player, or reintroducing a character who has temporarily moved to the background. Certain linguistic devices can be used to introduce new information, while other forms are reserved for situations when the speaker and listener share knowledge of the character (Brown & Yule, 1983; Levelt, 1989). The linguistic forms themselves are used as cues to signal to the listener how parts of the discourse are to be interpreted. Upon first mention, a character is unknown to the listener and, as such, should be introduced with an indefinite article (e.g., a frog), or a proper name.[1] These linguistic constructions indicate to the listener that the referent is new information. On the other hand, the identity of previously introduced story characters becomes shared knowledge between the speaker and the listener, which allows the speaker to use cohesive devices such as definite determiners (e.g., the frog), pronouns (e.g., *he*) or even zero anaphora (ellipsis) to refer to them. The speaker's choice of referring expression helps to guide the listener in interpreting the dialogue, while cohesive referential devices also help to bind the text, reduce redundancy, and abbreviate the referring expression.

The linguistic devices that create connections throughout a text are one way for the speaker to create cohesion. Cohesive elements are viewed by Halliday and Hasan (1976) as essential for connecting a series of sentences or utterances, thereby creating a unified whole and giving it the property of a text. A cohesive tie occurs between two items within a text: the item that specifies the identity, and the item that presupposes, or is dependent upon, the other item for its interpretation. In the example provided below, there is a cohesive tie between *he* and *a boy named Ryan*. The identity of the person being referred to by the pronoun *he* can only be recovered through its relationship with the noun phrase in the preceding sentence.

Once there was **a boy** named Ryan.

He had a dog, a frog, and a turtle.

[1] In some circumstances, the colloquial use of a demonstrative can also be acceptable for introducing a character (e.g., "one day *this* boy was going for a walk with his pets…"), particularly if a child is the narrator.

The relationship between the cohesive tie and its identifier can be either *anaphoric*, as in the example above, in which the character's identity *(boy*) is specified in an utterance preceding the cohesive element (*he*), or *cataphoric*, when the utterance specifying a character's identity follows the cohesive device later in the text. Moreover, these co-referential relationships can be expressed within a sentence or across several sentences. In the above example, the pronoun *he* creates a cohesive tie between the pronoun and the proper name in the preceding sentence, while *a boy* and *Ryan* is an example of intrasentential lexical cohesion. Within the system proposed by Halliday and Hasan (1976), the cohesive categories of interest for the present study include referential cohesion (i.e., personal and demonstrative), ellipsis, and lexical cohesion (e.g., repetition or naming).[2]

Personal reference includes personal pronouns (e.g., *he*, *she*, *they*, *it*) and possessive pronouns and determiners (e.g., *his*, *her*, *their*, *its*). Demonstrative reference consists of definite and demonstrative determiners (e.g., *the*, *this*, *these*). All of these forms are definite referring expressions, in that they assume that the identity of the intended referent is specifically known to the listener from a prior introduction. Definite forms are correctly used when they function to maintain the identity of a character that has been previously introduced with an indefinite determiner (i.e., *a*) or a proper name. When making the decision whether to use a nonspecific nominal form and an indefinite determiner (e.g., *a boy*) or a definite form (i.e., definite determiner, pronoun, or possessive form), the speaker must consider his own knowledge (i.e., do I have a specific entity in mind?) as well as that of the listener regarding the character being referred to. Many referential expressions contain little information about the character's identity;

[2] Halliday and Hasan (1976) define *referential cohesion* as one type of cohesive linguistic devices that are not interpreted on their own, but instead link to another mention of the person or object elsewhere in the text. Halliday and Hasan's definition includes only those forms that are cohesive. The definition of *reference* taken in this paper is more general and includes all mentions to story characters, regardless of whether they are cohesive or not; that is, they can stand on their own in the case of the introduction of a character.

this is particularly the case for pronouns, possessive forms, and ellipsis, which means that the listener must have a clear understanding of who is being talked about. If this is not the case, character identity might be ambiguous or erroneous ties may be made.

Two additional cohesive devices, ellipsis and lexical cohesion, can be used to refer to story characters, although these are far less common, especially in the narratives produced by children. Ellipsis (or zero anaphora) occurs when some part of the sentence (e.g., noun phrase or verb phrase) has been left unsaid, or omitted. Ellipsis can be used to refer to characters, but because the structure of English requires that all independent clauses (other than imperatives) include an explicitly stated subject, ellipsis is most commonly seen in conjoined clauses, as in the example below:

> The **frog** caught the bug and **Ø** ate him.

When the zero-form is used in a grammatically appropriate context, a cohesive tie has been created. In the above example, the identity of the character performing the second action must be recovered from earlier in the same utterance. Lexical cohesion creates a meaning relationship between two elements through the use of repetition or substitution. A frequent use of lexical cohesion is naming, as in the example presented above "a *boy* named *Ryan*". In addition, superordinate substitution can be seen in the following example:

> Once there was a kid named Ryan.
> He had a **dog**, a **frog**, and a **turtle**.
> He was walking to the city park.
> He was sure that his **pets** were having fun on the way.

Cohesive elements are essential for creating meaning relationships across the text, but their accurate use can prove challenging for the young language learner. In order to clearly refer to the characters in a story, the young narrator must first acquire the various linguistic forms used for referential cohesion. He must also find a balance between on the one hand providing enough information for the listener to clearly determine the identities of characters, and on the other hand utilizing cohesive ties to bind the text and reduce redundancy.

1.2. Acquisition of Referential Forms

Children acquire the grammatical forms needed to refer to story characters at a very young age, but the diversity and accuracy of use of such forms to introduce, maintain, and reintroduce the characters throughout a narrative continues to develop into the early school years. Developmental studies have focused on definite and indefinite determiners, pronouns, zero anaphora, and lexical cohesion.

1.2.1. Definite and indefinite determiners

Indefinite and definite determiners differ in their function: the indefinite form (*a*, *an*) is non-specific, referring to any entity in the class of objects specified by the noun; in contrast, the definite determiner (*the*) is specific to a unique object or individual. The indefinite form is appropriate for introduction of a character as the speaker must understand that, although the intended referent is known to him, the listener does not have any knowledge of this particular individual. Consequently, an indefinite noun phrase must be used. For subsequent mentions of the same character, the definite determiner (*the*) calls up the same unique referent for both the speaker and the listener. Making the appropriate selection between indefinite and definite noun phrases requires cognitive awareness on the speaker's part, as he must be aware of both his own and his listener's knowledge states.

Brown (1973) reported evidence from naturalistic observations that 3-year-old children were correctly using indefinite and definite determiners to talk about specific and non-specific objects, but that these forms were mostly used to name objects that were present in the physical context. Similarly, other research has suggested that when young children attempted to use these forms to introduce characters in narrative discourse, they initially relied on cues provided by the context or the physical situation to identify characters, often in the form of pointing to the characters in the pictures (Beliavsky, 2003; Hickman, 1980; Karmiloff-Smith, 1980). Hence, preschool children used indefinite and definite forms in qualitatively different ways from older children, often relying on extra-linguistic information (i.e., labeling what was seen in the pictures) as opposed to using linguistic devices to specifically refer to characters within a narrative framework.

Research has shown that as children begin to develop the ability to use textual reference, references to story characters become embedded in the text, with linguistic devices specifying the identity of the characters. Children also progressively become more accurate in their use of these forms, which presumably corresponds to a greater ability to consider not only their own knowledge, but also the listener's familiarity with the intended referent.

The developmental literature regarding children's acquisition of the indefinite/definite distinction for determiners is not consistent, with research reporting acquisition over a broad developmental period (i.e., between 3 and 9 years of age). Warden (1976) found that children under the age of five used predominantly definite expressions to introduce a referent, which suggests an inability to take into account the knowledge state of the listener. In contrast, the groups of 5- and 7-year-old children in the same study could accurately use the indefinite form for introduction 62% and 61% of the time respectively. In addition, the narratives produced by these age groups included several instances of self-correction, which often signals an emerging skill. Only the 9-year-olds in Warden's study reliably introduced characters with indefinite determiners in a story telling task.

Hickman (1980) also examined children's use of referential devices for introduction and maintenance in their narrative productions elicited from short films. Results of this study were consistent with Warden's findings, with 7-year-olds using definite forms (i.e., definite article, pronoun) almost as frequently as indefinite forms (i.e., indefinite article, demonstrative) to introduce new referents. The appropriate distinction between indefinite/definite articles appeared in the narratives of the 10-year-old children. A high level of success was seen for maintaining reference to story characters across ages, with 5-, 7-, and 9-year-olds accurately using the definite form for this function (Hickman, 1980; Warden, 1976).

In contrast to these findings, early work by Maratsos (1974) found evidence of accuracy with the indefinite and definite forms at an earlier age. The study by Maratsos used two different tasks. First, in the elicitation task, the experimenter told the participant a story, followed by questions about the characters. The stories and questions were designed to elicit either an indefinite or a definite noun phrase. Second, in the imitation task, the child repeated back sentence by sentence a story told by the experimenter. The experimenter intentionally omitted all the articles, and examined whether the child inserted an indefinite or a definite determiner. Results indicated that 3- and 4-year-old children had a rudimentary understanding of the indefinite/definite distinction, as both age groups used the appropriate form in the elicitation task at a level of performance above chance. With the imitation task, 94% of the articles supplied by the children were accurate. Additionally, Maratsos found that the 3-year-old children relied more heavily on indefinite rather than definite forms when referring to story characters, which contrasts with the preference for definite forms reported by both Hickman and Warden. Although the 3 year-olds may have been relatively more successful with the indefinite forms, Maratsos did not interpret this finding as evidence that these young preschool children understood the knowledge state of their listener; rather, the author proposed that the demands of the tasks were too high, and thus did not allow for the children to establish whether a referent was unique and specified for themselves or for the listener. The data for the 4-year-old children in the Maratsos study was more consistent with the results of Hickman and Warden, as this age group was highly accurate (95%) with the definite forms, but still made many errors when the indefinite form was required (79% correct). Based on the data from her study, Hickman concluded that the metalinguistic skills needed for accurate use of definite and indefinite determiners emerged around the age of 7. This conclusion is supported by the attempts at self-correction made by the children who participated in Warden's research. Selecting the correct form to introduce a character is dependent on several skills. First, the child must be aware that the different referential forms serve different functions and that there is a need to clearly identify characters prior to using specific forms to refer to them. Second, reference within a narrative framework requires the social cognition to manage differing speaker and listener knowledge states. As a whole, the evidence suggests that, although younger children may have a basic understanding of the different function of the indefinite and definite determiners, accurately managing these forms in a complex narrative task is a skill not acquired until around the age of 7 years.

1.2.2. Pronouns

It has been well established in the literature that, prior to the age of 5, children cannot manage pronominal forms in narratives of significant length or complexity (Bamberg, 1987; Bennett-Kastor, 1983; Karmiloff-Smith, 1980; Wong & Johnston, 2004). Because pronouns necessarily stand for something or someone, they require inference to determine the identity of the referent, which makes them more demanding for both the speaker and the listener. Developmental data on the use of pronouns has indicated that between the ages of 4 and 5, pronouns are not used unambiguously and require some degree of extralinguistic information for interpretation (i.e., spatial location, pointing, or shared knowledge; Karmiloff-Smith, 1980; 1985). From this data, Karmiloff-Smith created a three-stage model for the acquisition of pronominal forms. In stage one, children used nominal and pronominal forms interchangeably, not understanding the cohesive or anaphoric property of pronouns. In the next stage of development, which emerged around 6 years of age, children were reported to show a *thematic subject strategy* in which pronominalization was reserved for the main protagonist. For Karmiloff-Smith, pronoun use at this stage reflected a way for the children to organize their narratives, as opposed to serving an anaphoric or an economizing function. In the final stage, which occurred predominantly in 8- and 9-year-old children, pronouns were used to refer to secondary characters as well as primary characters, but this was reported to be rare and to occur primarily within a single sentence following a conjunction.

In a large cross-sectional study, Bamberg (1987) elicited narratives from German-speaking children aged 3 to 9 years using a wordless picture book. The focus of Bamberg's study was to describe the distribution of the different types of referring expressions (e.g., noun phrases, pronouns) by their discourse function. Two referential functions were considered, switching and maintenance. A referring expression was classified as switching if it was the first mention of a story character (introduction) or if the narrator returned to a previously mentioned character (reintroduction). Maintenance occurred when the narrator continued to refer to the same character over successive utterances. The accuracy of pronominal forms (pronouns and ellipsis) was evaluated depending on *referential function*. Results showed that 3- and 4-year-old children used pronouns for both switching and maintaining reference to characters. This trend of overusing pronominal forms for all referential functions decreased with age, while the tendency to use pronouns to maintain reference to characters increased with age. Therefore, preschool-age children had not yet distinguished the usage differences of noun phrases and pronouns depending on referential function, and only the 9- and 10-year olds consistently reserved pronouns for maintenance. When compared to the findings of Karmiloff-Smith (1980), although there was agreement across studies regarding the age of acquisition of the referential devices, Bamberg's data showed inconsistent evidence of a thematic subject strategy. The narratives produced by the 3- and 4-year-olds did show some evidence of a thematic subject strategy, but this was not the case for the 5- to-6-year-old children, as the distribution of pronoun forms was relatively equal across the two characters in the story (Bamberg, 1987).

The early work by Bamberg (1987) and Karmiloff-Smith (1980; 1985) reported on the distribution of referential devices of various types according their function in the story, but it was not until later research that the accuracy of the pronominal forms was evaluated. Roth, Spekman, and Fye (1995) reported very successful use of referential devices by 8- to 10-year old children: referential adequacy was on average 86% correct in narratives elicited from a picture and 96% correct in spontaneous narratives. Furthermore, accuracy reached 93% for articles, indicating differentiation between the definite and indefinite forms by 8-year-old participants. These findings are in accord with the results of other studies. Liles (1985b) reported that 7- to10-year-old children used adequate cohesive ties 86% of the time in the retell of a film story, and Finestack, Fey, & Catts (2006), found that second grade children were accurate for 84% of referential ties in a picture-elicited narrative. In a study by Beliavsky (2003), the children previewed a wordless picture book before producing their narratives in a manner that limited the shared context between the speaker and the listener. Results indicated that for kindergarteners, 16% of the pronouns used involved *exophoric* reference (i.e., using the extralinguistic context), which drastically reduced to only 1% in grade 1 and 2 children. Correspondingly, cohesive anaphoric reference (i.e., made explicit within the text) increased with age, with considerable improvement occurring from kindergarten (63% accurate) to grade 1 (90% accurate), but relative stability between grades 1 and 4. In conclusion, the results presented above suggest that by the age of 7 years, children are fairly successful at clearly referring to characters within narrative discourse. Character maintenance, primarily through the use of pronouns, shows greater accuracy at an earlier age; in contrast, introduction and reintroduction, which rely heavily upon the mastery of indefinite and definite forms, continue to develop until the age of nine.

1.2.3. Zero anaphora and lexical cohesion

Developmental studies on the use of zero anaphora and lexical cohesion have indicated that these forms were also present at an early age. Berman and Slobin (1994) reported the use of null subjects in their cross-sectional data of narratives elicited using Mercer Mayer's (1969) wordless picture book *Frog Where Are You?*. Data from the English-speaking children showed that the youngest age group of 3- and 4-year-old children were using zero anaphora in two different situations: the first resulted in an ungrammatical production, with subject ellipsis in an independent clause (e.g., Ø go in there); the second occurred in response to questions occurring in conversation, which was considered accurate and appropriate. In the older group, 5- and 9-year-old children used null subjects in coordinated clauses (e.g., **He** climbed on a big rock and **Ø** called out for him). These accurate productions occurred with and without markers of coordination.

The use of lexical cohesion was examined by Liles (1985b) in a study that compared reference to characters in narratives produced by children with and without language disorders. The narratives were produced as a retell following the viewing of a film. Results showed that the children with language disorders, aged 7-10 years, relied more heavily upon demonstrative reference and lexical cohesion when referring to characters compared to the typically-achieving children. Liles interpreted the use of lexical cohesion by the children with language disorders as corresponding to "simple descriptions of what was seen rather than intended as cohesive explanations of the relationships between characters and events" (p. 418). Furthermore, she suggested that children may rely more heavily on lexical cohesion when they have not yet developed accurate use of the pronominal system. With lexical cohesion providing much more information as to the identity of the characters being referred to, it may offer a less complex way for creating cohesion, whereas pronouns and ellipsis could require a higher level of language skill in order to be used accurately in narrative discourse.

1.3. Contextual Factors in Narrative Productions

The elicitation context has been found to influence how the narrator refers to story characters. The influence of the testing environment on narrative production was apparent in a study by Wigglesworth (1990) where two sets of picture were used to elicit narratives from children aged 4, 6, and 8 years, as well as adults. The first set of pictures was presented to the participant without indication that they comprised a story, and the participant was asked "what's happening?". These instructions led the children to interpret the context as a descriptive task, as opposed to a narrative one. Additionally, the experimenter was seated next to the participant, which created a situation of shared knowledge. Such contextual factors likely accounted for the inappropriate use of definite forms and of ellipsis to introduce characters by all three groups of children as well as by the adult control group. Warden's 1976 study (see above) also revealed the significant effect that shared context can have on a speaker's choice of linguistic forms. Experimental conditions in which the listener viewed the story pictures alongside the child resulted in productions that contained more extralinguistic and definite forms to introduce characters. Furthermore, under these shared-context conditions, even the adult comparison group did not consistently use indefinite expressions when initially speaking about the characters. In contrast, when the shared knowledge between the speaker and the listener was eliminated by introducing a *'naïve'* listener, all of the adults used indefinite forms for the first mention of story characters and objects, whereas the 7- and 9-year-olds did so 61% and 82% of the time respectively. Results from Liles (1985a) were consistent with those of Warden; when the listener did not view the film used to elicit the narrative, the children produced longer stories with a higher proportion of cohesive referential ties. Such results indicate that children are sensitive to the needs of their listener: when the listener is already familiar with the story, children are more likely to use linguistic forms that reflect the shared knowledge of the identity of the story characters.

Context is not the only factor that has been found to impact children's performance with referential devices, as great consideration also needs to be given to the demands of the elicitation task. For instance, Wigglesworth (1990) reported very little evidence to support the presence of a thematic subject strategy in the stories of 6- or 8-year-olds. She speculated that this may have been due to the active participation of the secondary characters in the pictures used to elicit the narrative. This explanation is supported by the work of Bamberg (1987), which also failed to provide evidence for a thematic subject strategy (see above). In this case, the wordless picture book that was used to elicit the narratives contained two characters that were central to the plot. Taken together these studies suggest that when there was no clear delineation between a single main protagonist and other minor background characters, the pattern described by Karmiloff-Smith did not emerge.

When trying to determine a child's true referential abilities, it appears crucial to judiciously select the elicitation task while carefully controlling the environment. Emslie and Stevenson (1980) used very simple stories consisting of three pictures to elicit narratives from their participants. This resulted in the adequate use of referential forms by 3- and 4-year-old children—a substantially younger age than many other experimental studies have found (Bamberg, 1987; Beliavsky, 2003; Karmiloff-Smith, 1980; Warden, 1976). Emslie and

Stevenson suggested that if task difficulty was matched to the child's cognitive abilities, children would be much more successful in producing clear referential ties, and would be able to do so at an earlier age. However, the complexity of the narrative task must also be sufficient enough to challenge the participants if one hopes to reveal individual differences and developmental trends.

1.4. Using Referential Forms in Narratives: Discourse Models

In order to accurately refer to story characters throughout a narrative production, a child must have acquired the requisite lexical forms. This is not sufficient, however, because creating these meaning relationships within and across sentences while simultaneously adapting to the listener's needs requires significant cognitive resources and abilities. Complete and unambiguous reference to story characters will largely depend on the speaker's ability to change perspectives. Cornish (1999) described the speaker's role as one of choosing a referential device that coincides with the level of cognitive accessibility of the intended referent in the listener's knowledge (or memory) of the story. As such, if the character does not exist in the listener's knowledge, the speaker must introduce it; conversely, if the speaker can assume that the listener already has a representation of the intended referent, then he may select a pronominal form that is consistent with the listener's focus and attention to that character.

A discourse representation or discourse model is a specific mental representation of an individual's knowledge of a discourse experience. It is constructed based on what has occurred in the discourse thus far, combined with one's perception of the speaking context and more general representation of the world (Brown, & Yule, 1983; Johnson-Laird & Garnham, 1980). In the case of narrative discourse, the speaker has a particular state of affairs in mind and produces a narrative based on this representation. Upon hearing the narrative, the listener also progressively builds a discourse model of the narrative based on how it has been communicated by the speaker. In many cases, the narrative elicitation context used in research involving children inherently produces a disparity between the discourse models of the speaker and the listener. For instance, if the child is provided with a wordless picture book to look through and then asked to produce a corresponding story to a naïve listener, the speaker has prior knowledge of the story that is not shared with the listener. Consequently, the speaker must understand that the listener is not familiar with the pictures or the story and formulate his utterances accordingly. A discourse representation is not static, as it is continually changing with each utterance as the story unfolds. It is via the discourse model that the speaker tracks what he has already said, what he intends to say next, and what referential devices would be appropriate given the knowledge he presumes the listener to have. Several theories have been proposed to explain how speakers create a discourse model.

The speaker's discourse model includes the characters and the events of the narrative as well as what he presumes about the knowledge of the listener, which together will guide the speaker in selecting the appropriate linguistic forms to refer to characters. Brown and Yule (1983) described a referent as either *given* or *new* information for the listener. If the character has yet to be introduced and is thereby unknown to the listener, it represents *new* information. In contrast, information that is *given* is that which the listener already knows, as it has been previously introduced and thus should be recoverable from the text. The dichotomy of *new* and *given* roughly corresponds to the use of indefinite and definite forms, but further distinction is required to explain the distribution of definite forms in English. Brown proposed two additional subcategories of *given* information which depend on how recently the reference to a character has been made. *Current* given information is that which is in the consciousness of the listener at that particular moment, and allows for pronouns and ellipsis to be used as referential devices. *Displaced* information, while familiar to the listener, has not been mentioned in the immediately preceding utterances; as a result of this greater distance, more descriptive referential forms (i.e., a definite article and a noun) must be used to direct the listener to switch his attention to the correct referent.

The Givenness Hierarchy (Gundel, Hedberg & Zacharski, 1993) was specifically developed to explain how a speaker uses a discourse model to select appropriate referring expressions. It is similar to Brown and Yule's model, with listener familiarity forming the basis for selecting referential devices. The Givenness Hierarchy is a continuum of six cognitive statuses that represent how familiar the listener is with the speaker's intended referent:

type identifiable < referential < uniquely identifiable < familiar < activated < in focus

Each cognitive status is associated with certain referential devices. At one end of the continuum is the *type identifiable* status, which represents the lowest level of listener familiarity. This status does not assume any previous knowledge on behalf of the listener, but does assume that the listener knows which class of object the noun refers to. *Type identifiable* is the cognitive status that corresponds to the use of indefinite forms (*a*/*an* + noun). The other end of the continuum is anchored by the *in focus* status, which represents the highest level of listener familiarity. This status represents the objects and characters that are currently the centre of attention in the narrative and thus in the listener's focus of attention; consequently zero anaphora and pronouns can be used by the speaker to refer to them. Each cognitive status along the continuum assumes an increasing level of knowledge and attention of behalf of the listener, and the corresponding linguistic forms that are associated with each cognitive status vary in the amount of knowledge that they presuppose. For instance, all pronouns require that the referent be activated, as the minimal descriptive content of pronouns offers little basis for identifying the referent. For a referential form to be used appropriately, it must meet the requirements of the corresponding cognitive status. Referential forms corresponding to lower statuses can be used, since each higher status of the Givenness Hierarchy assumes all lower statuses. Nonetheless, according to Grice's maxim of quantity (Grice, 1975), the speaker should attempt to provide only what is required and nothing more. This principle would be violated if the speaker used forms that maximally identified the referent in all instances. Indefinite forms will always meet the necessary conditions for all cognitive statuses, since they are the least restrictive and accordingly appropriate for any referent. However, relying only on indefinite forms would not respect Grice's maxim. Gundel and colleagues (1993) proposed that it is the interaction between the Givenness Hierarchy and Grice's maxim of quantity that determines first, how much a speaker can presuppose based on the listener's attention and activation levels and, subsequently, which linguistic forms can be appropriately selected for complete and unambiguous reference.

The discourse model is constructed in real time by both the speaker and the listener as the narrative unfolds and it becomes increasingly more complex. When constructing a narrative of any length, the speaker cannot keep the whole discourse model within his centre of attention. As such, Levelt (1989) proposed that the discourse model is stored in long-term memory, with only a small fragment in focus at one time. The same holds for the listener. Information that the speaker believes both he and the listener are attending to is considered in focus and is currently activated in memory. Information that has been introduced but is not currently active in the story development remains in the discourse model, but not in focus. As the speaker moves between characters when creating the story, he must continually change what is held in working memory, bringing to the forefront previously mentioned information from long-term memory and holding it active in working memory. The speaker then has to evaluate the familiarity of the intended referent to select the appropriate referential forms. It is through the use of a discourse model that the young language learner can track the progress of his narrative, while simultaneously monitoring the knowledge and awareness of his listener. The constant exchange of information and revision of the discourse model is presumed to be the function of the updating process.

1.5. Cognitive Aspects of Narrative: Working Memory and Updating

Working memory is a limited capacity system responsible for online monitoring, processing, and maintenance of information (Baddeley, & Logie, 1999). The most influential model of working memory is the well-studied multi-component working memory model first introduced by Baddeley and Hitch in 1974, and later modified to include a fourth component (Baddeley, 2000; Baddeley, 2003a; Baddeley, & Logie, 1999). This model includes; the *phonological loop*, a temporary store for verbal-acoustic information; the *visuospatial sketchpad*, which stores and processes spatial, visual, and kinesthetic information; the *episodic buffer*; and the *central executive*. The central executive is a supervisory system of limited attentional capacity that controls and coordinates the working memory system. The episodic buffer is the most recent addition to the model, added specifically to combine information from the other subsystems and from long-term memory. The episodic buffer is “assumed to be attentionally controlled by the executive” (Baddeley, 2003b, p. 836). The central executive’s role in the mental manipulation of material held in working memory seems particularly relevant when considering the demands of online narrative production, and in particular the challenge of clearly referring to story characters.

Several research tasks have been designed to test the capacities of each of the three components proposed by Baddeley. Short-term memory (STM) tasks are designed to measure the immediate recall of verbal or visuospatial information. Common measures of verbal STM include the digit span and the word recall task, in which the participant hears a series of numbers or words and has to recall the sequence in the same order. Two commonly used visuospatial

STM tasks include the dot matrix task and the visual immediate memory task, both of which require the participant to recall the location of dots or coloured squares in a grid. Measurement of the central executive began with tasks designed to measure both the storage and processing functions of the memory system. The most widely used working memory tasks are the backwards digit recall and variations of Daneman and Carpenter's (1980) reading span. The reading span requires the participants to read aloud a series of sentences while remembering the final word in each of the sentences. Adaptations to Daneman and Carpenter's original procedure have lead to various verbal and visuospatial complex span tasks. The processing requirements were increased in a task used with children by having the participant judge the validity of the sentences as they listened to them, while also remembering the final words (Alloway, Gathercole, & Pickering, 2006). More recently, researchers have worked to further specify the role of the central executive by examining how it dynamically changes the contents held in working memory, a process labeled updating.

Updating and monitoring the contents of working memory (henceforth, *updating*) is an executive function responsible for encoding incoming information relevant for the ongoing task, and then, as appropriate, revising the contents held in working memory, by either replacing or modifying the old information with newer, more relevant, information (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000; Morris & Jones, 1990). Kessler and Meiran (2008) elaborated on this definition by identifying that there is both the need to modify the contents as well as to maintain unchanged the still relevant information held in working memory. When incorporated into Baddeley's working memory model, the updating function would reside in the central executive. Tasks have been designed to measure updating in both the visuospatial and verbal domains. The visual n-back task requires participants to compare the image they are currently viewing to images presented one or two items previously. The keep track is a verbal task that requires the child to remember the last item presented in several target categories as a series of items is presented. The critical feature that distinguishes an updating task from a complex working memory span task is that, in the case of updating, active manipulation of the contents in working memory involves replacing old information with new information. Although much of the more recent research on the executive functions has highlighted the difference between a complex working memory span task and an updating task, strong correlations have been found between two types of tasks, leaving some researchers hesitant to conclude that they correspond to separate processes (Lehto, 1996; Schmiedek, Hildebrandt, Lovden, Wilhelm, & Lindenberger, 2009; St. Claire-Thompson, & Gathercole, 2006).

Baddeley (1996a) conceptualized the central executive as a unitary, domain-general system for the control and manipulation of information in memory. This entails that both complex working memory span tasks and updating tasks, either in the verbal or the visuospatial domain, rely on the same limited executive resources (Baddeley, 1996a, Baddeley, 1996b). Some experimental studies have supported the unitary central executive proposed by Baddeley and colleagues, while others have provided evidence suggestive of a fractionated central executive.

Proponents of the domain-general view of the central executive have cited evidence of high correlations between performance on complex visuospatial working memory span tasks and complex verbal working memory span tasks, suggesting that these tasks are measuring the same construct (Alloway et al., 2006, Engle, Kane, & Tuholski, 1999). The evidence provided by Alloway and colleagues (2006) comes from a large scale study of children ages 4 to 11 years that used a confirmatory factor analysis. The model providing the best fit of the data included two domain-specific STM systems, one for visuospatial information and one for verbal information, along with a domain-general executive control centre. This evidence suggests that the complex working memory tasks called upon the short-term systems for the storage of information, and on the general central executive for the processing and manipulation of the information.

In contrast to evidence for a three-factor model, other research has found a strong association between visuospatial STM tasks and visuospatial tasks measuring the central executive, including both complex working memory spans and updating tasks (Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M., 2001; Morris, & Jones, 1990; Yue, Zhang, & Zhou, 2008). Results from these studies suggest that the visuospatial sketchpad responsible for short-term storage and rehearsal of visual or spatial information may to be more closely related to the central executive. Miyake and colleagues (2001) argued that the storage and rehearsal of verbal information is much more practiced compared to performing such tasks in the visual or spatial domains. Consequently, short-term maintenance of visual or spatial information may be more dependent on the resources of the central executive. Contrary to the results of Alloway et al. (2006), Shah and Miyake (1996) found that, when testing adults, performance on complex visuospatial working memory tasks did not predict performance on complex verbal working memory tasks. Such findings suggest that the two tasks were not drawing upon the same domain-general central executive resources. Furthermore, performance on the complex verbal working memory task (reading span) successfully predicted performance on a reading comprehension test, whereas performance on the complex visuospatial working memory measure (spatial span) did not.

Several studies have furthered our understanding of updating by manipulating the demands placed on the central executive and measuring the effects this had on task accuracy. Kessler and Meiran (2008) found that the total amount of information held in working memory affected the ability of participants to effectively update memory contents. Additionally, updating more than one item at a time or updating the contents held in memory numerous times both resulted in degraded performance (Carretti, Cornoldi, De Beni, & Romano, 2005; Yue et al., 2008). Finally, the level of activation in working memory affected how easily a participant could update and later recall a remembered item. If an item was relevant for some time before it had to be replaced, it was more likely to result in an intrusion error when compared to an item that could be discarded almost immediately (Carretti et al., 2005).

Several studies have identified updating as a distinct executive function that is important for many higher order cognitive tasks. Nonetheless, research thus far is not as clear as to whether updating is best conceptualized as a single, domain-general resource, or whether separate processes are responsible for accomplishing verbal and visuospatial updating tasks. Additionally, although the distinction between working memory and updating has been well-defined theoretically, results have shown a strong correlation between these two components of the memory system. These strong correlations present the possibility that complex span tasks and updating tasks are in fact measuring the same underlying process.

1.6. The Relationship Between Language, Cognition, and Updating in Children

A relationship between updating and language ability has been identified in research examining reading, verbal reasoning, and language comprehension in children. Carretti, Cornoldi, De Beni, & Romano (2005) found that reading comprehension and the ability to update the contents of working memory were related in a study that included children aged 8 to 11 years. The participants were divided into two groups based on scores on a standardized reading comprehension test; *poor comprehenders* scored below the 25th percentile, whereas *good comprehenders* obtained scores above the 75th percentile. In the updating task, the child viewed a vertical column of pictures while listening to the experimenter read a list of words. The child was to remember the three words that corresponded to the pictures that were closest to the top of the picture column. Results showed that the good comprehenders were able to correctly recall more items than the poor comprehenders. This study also discussed the important role that inhibition plays in working memory. Updating the contents of working memory requires that all or part of what was previously being held on to must be discarded; in other words, the information that is no longer relevant must be inhibited. The authors concluded that difficulties inhibiting previously relevant information that became irrelevant as the task proceeded might have contributed to poor updating performance by the group of poor comprehenders. Therefore, the ability to suppress irrelevant information, through the use of inhibition, may be an important factor for successful performance on an updating task.

In another study, the relationship between updating and higher order cognitive abilities was evaluated in grade 4 and 5 children (van der Sluis, de Jong, & van der Leij, 2007). Three different verbal updating tasks were used in this study including a version of the keep track task, as well as letter memory and digit memory. In letter memory, the participant was presented with a series of written letters and asked to remember the last three presented. The number of letters in the series was random, so the participant had to continuously update the to be remembered information as each item was presented, not knowing when the series would end. Number memory was a numerical version of letter memory. The results showed that updating ability was positively correlated with scores on standardized verbal and non-verbal reasoning, reading, and arithmetic tests. Although positive correlations were found, the sizes of the correlations were small to moderate, explaining only 2.6-15.1% of the variability across the four abilities measured. The authors highlighted the difficulty of task impurity in measuring executive functions, and that this likely contributed to the low correlations found. Executive functions by definition are the control mechanisms of other cognitive processes. As such, tasks that tap executive function always involve other cognitive abilities and processes, including verbal abilities, speed of processing, visual-spatial skills, or storage capacity, to name a few. Although the use of factor analysis and the inclusion of control variables have resulted in the identification of updating as a distinguishable executive function, the task impurity problem has made it difficult to distinguish which cognitive and executive processes contributed to performance on the complex language tasks (Miyake et al., 2000; van der Sluis et al., 2007).

A recent study looked at the updating abilities of children aged 7 to 12 years with and without language impairments (Im-Bolter, Johnson, & Pascual-Leone, 2006). In this study the children completed the visual n-back task. In this updating task the participant had to determine if the pattern of dots they were looking at was the same or different as the pattern just previously presented. The group of children with language impairments obtained lower scores on the visual updating task when compared to the group of typically-achieving children. Furthermore, the results suggested that updating mediated the relationship between domain-general mental attention and language competence. The relationship between updating and language competence revealed in this study suggests that updating is a domain-general ability, given that even when updating was measured in a visual task that invited little verbal mediation, a significant relationship with language abilities nonetheless emerged.

The research on updating skills in children has found that updating is a factor related to performance on language comprehension, reading, and other higher-level cognitive tasks. Work by Im-Bolter and colleagues (2006) presents some evidence for a domain-general updating process, but research has yet to examine both verbal and visuospatial updating tasks with the same group of children. Additionally, research has yet to explore the development of updating skills of young children.

1.7. The Present Study: Updating and Referential Adequacy in Children's Narratives

Clearly referring to story characters is not a simple task, particularly for a young child. It requires having acquired the various referential forms, as well as understanding the functional distinctions between each linguistic device, and applying this knowledge online during the narrative production. Successful narration requires the manipulation of a dynamic discourse representation, which then allows the speaker to create accurate presuppositions, and thus accurate reference. Throughout the generation of a narrative, the speaker must utilize working memory resources to update the discourse model, adding new information and reactivating previously mentioned characters as the story progresses. The ability to update working memory representations is suspected to be a requisite skill for adequate maintenance of the discourse model and the selection of referential forms.

This study investigated whether individual differences in updating and adequate reference to story characters in narrative discourse were related in a group of children from kindergarten through second grade. It also documented developmental trends in both of these abilities, and examined factors that may have contributed to the difficulty of clearly referring to story characters. To serve these goals, this study attempted to answer the following specific questions.

Research Questions:

1. Is children's ability to adequately refer to story characters related to their ability to update the contents of working memory?

   Although this specific question has not yet been addressed in the literature, a significant positive relationship was expected between updating and referential adequacy based on theoretical grounds and on related research that has highlighted a relationship between language abilities and executive functions.

2. Are there significant developmental trends evident in children's updating ability?

   Although very little research has looked at the development of updating capacity in young children, related research in executive functions and brain development (e.g., Diamond, 2006; Hitch, 2006) suggested that children between kindergarten and grade 2 show an increase in performance on updating tasks with age. A description of the developmental progression of these skills will represent an additional important contribution to the existing literature.

3. Are there significant developmental trends evident in children's ability to adequately refer to characters?

   Based on prior research, it was expected that the older children would be more successful than the younger children in adequately referring to characters. As such, referential adequacy was expected to be positively and significantly related to age. In addition, the diversity of forms used to refer to characters was anticipated to increase with age. In particular, reliance on nominal forms (i.e., lexical cohesion) was expected to decrease in favour of greater use of pronominal forms (i.e., referential cohesion). Finally, it was hypothesized that the developmental trends in referential adequacy could vary depending on referential type (in particular for nominal compared to pronominal forms).

4. Is there evidence of a stage of development in which children rely on the thematic subject strategy for reference to story characters?

   Many research studies have failed to replicate Karmiloff-Smith's findings of a thematic-subject strategy used by school-aged children. It was not expected that the children in this study would show a preference for using pronouns exclusively for maintaining reference to the main protagonist while utilizing noun phrases for maintenance of secondary characters.

5. Does the ability to clearly refer to characters vary systematically depending on referential function?

   Based on previous research, it was expected that the children would achieve the highest level of referential adequacy when maintaining reference to story characters, as this function was expected to be the least demanding on memory processes, and the linguistic forms for maintenance are acquired at a young age. It was expected that the younger children would exhibit some difficulties with introduction because they would still be acquiring the appropriate use of the indefinite form. Reintroduction was hypothesized to be most demanding on working memory resources and, as such, was expected to show the strongest correlation with updating scores and the lowest rates of referential adequacy.

2. METHOD

To investigate the relationship between the ability to update the contents of working memory and the ability to clearly refer to characters in a narrative, the participants completed two story-telling tasks and three measures of updating. The three updating tasks were chosen to investigate both domain-general and domain-specific resources in updating abilities, with the individual tasks making more or fewer demands on the auditory-verbal and the visual components of working memory. Measures of visual and verbal short-term memory were also obtained to investigate whether storage capacity mediated the ability to update the contents held in working memory. This chapter describes the characteristics of the study participants, along with the experimental and scoring procedures.

2.1. Participants

Thirty-nine typically-developing, monolingual English-speaking children between the ages of 5 and 8 years, from kindergarten through second grade, participated in the study. Participants were recruited from two schools in the Maple Ridge School District, a suburb near Vancouver, Canada. Participants meeting the inclusion criteria were identified by the classroom teacher, following which consent was obtained from the students' parents. Monolingual English speaker was defined as an individual who spoke primarily English both in the home and at school with minimal exposure to any other language. Participants were deemed typically developing if there was no report of language or learning concerns. Information corresponding to these inclusion criteria was confirmed through teacher report and a parent questionnaire. Of the 39 participants who were initially recruited, two kindergarten children declined to complete the testing and were excluded from the analysis. Altogether, this study includes data for 37 children, 12 children in kindergarten (5 boys, 7 girls; mean age, 5.89 years), 13 in grade 1 (7 boys, 6 girls;

M age, 6.87 years), and 12 in grade 2 (5 boys, 7 girls; *M* age, 8.09 years). Approximately equal numbers of girls and boys participated in each grade, 17 boys and 20 girls in total. The children came from varied socio-economic backgrounds, with four mothers (11%) reporting high school education, 23 (62%) having received some additional schooling, and 10 (27%) being university-educated. The children were well matched across grades regarding the level of maternal education as measured in the number of years of schooling. See Table 1 for additional details.

Table 1: *Demographic data (means and ranges), by grade*

	Grade		
	K (*n* = 12)	1 (*n* = 13)	2 (*n* = 12)
Age (yrs)	5.89 (5.50-6.33)	6.87 (6.42-7.25)	8.09 (7.58-8.42)
Maternal educations (yrs)	14.3 (11-18)	13.5 (12-17)	14.3 (12-17)

2.2. Experimental Tasks

2.2.1. Procedures

All the participants completed seven tasks in total: two tapping short-term memory storage, three assessing the ability to update working memory, and two narrative productions from wordless picture books. Testing took place in two sessions held on separate days, with participants seen individually. All testing was completed during the last two months of the school year.

Testing included two measures of short-term memory capacity. The *visual span task* was modeled after the recognition version of the visual immediate memory task designed by Logie and Pearson (1997). This task was selected because it is visual, as opposed to spatial, and it is a recognition task, which matched the design of the n-back procedure. A matrix of squares, half of which were coloured red and half of which were blank, was presented on a computer screen. The matrix was displayed for 2 seconds, followed by 2 seconds of blank screen, after which the matrix reappeared with one of the previously red squares now blank (see Figure 1 for an example). The distribution of the red and blank squares, as well as the square that changed from red to blank, were determined at random. The child was instructed to point to the square that had changed from red to blank. The response matrix remained on the screen until the child indicated his response by pointing to a square on the screen. The child's response was recorded online by the experimenter; the task was also video-recorded to provide a back-up to verify scoring accuracy.

The task began with each participant completing one practice trial consisting of a 2 x 2 matrix, followed by three test trials at this level. If the participant answered correctly on at least two of the three trials at any given level, the size of the matrix increased by two squares (one red, one blank), with the squares progressively added to start and then complete a new row and then a new column. All children received identical trials in the same order. For each level, the same discontinue rule applied: the task ended when the child responded correctly on fewer than two of three trials. The highest possible level consisted of 5 x 6 matrices, with 15 red squares and 15 blank squares. Each participant's span was calculated based on the mean number of red squares in the three largest matrices in which the child responded correctly.

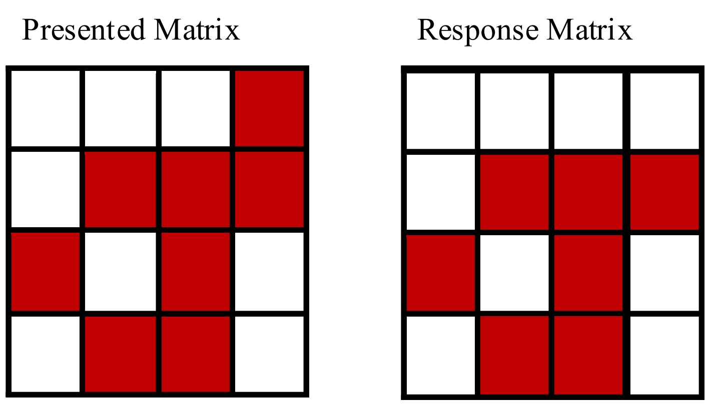

Figure 1: Example of a trial from the visual span task

Verbal short term memory capacity was assessed via a *digit span task*, the Number Repetition-Forward task from the Clinical Evaluation of Language Fundamentals-Fourth Edition (CELF-4; Semel, Wiig, & Secord, 2003). This task was selected because it is a widely used measure of verbal storage capacity, and it requires a verbal recall response, which is consistent with the keep track task. The digit span was administered according to the standardized procedure. The experimenter recited a string of digits, and the child was told to repeat the sequence in the same order in which he/she had heard it. Two trials were presented at each level, with the first level consisting of strings of two digits. If the child correctly recalled all the digits in the correct order for at least one of the trials within a level, the task progressed to the next level with trial length increasing by one digit. The task was discontinued when the child failed both trials of a given length. The highest possible level consisted of sequences of nine digits. The participant's score was calculated by adding up the number of trials recalled correctly.

Three tasks were selected to evaluate the ability of participants to update the contents held in working memory: *n-back*, *sound monitoring,* and *keep track*. These tasks were chosen specifically because they varied in terms of the nature of the to-be remembered material (visual, auditory, or verbal), the response required (non-verbal or verbal), and the extent to which verbal mediation was likely to support performance. All of the tasks were modeled after updating procedures used in previous research with adults and children, and each was adapted to be appropriate for use with young children.

The *n-back* task (Im-Bolter et al., 2006; Mantyla, Carelli, & Forman, 2007) was used to measure updating in the visual domain; it involves visual presentation, requires a nonverbal response, and offers little potential for verbal mediation. The 1-back and 2-back procedures used in this study were programmed and administered using E-Prime software (Schneider, Eschman, & Zuccolotto, 2002). Various configurations of three white dots on a black background were presented in the middle of a computer screen, one at a time. Each image was displayed for 2000 ms with 500 ms of blank screen before the next image appeared. The child was instructed to press the spacebar if the configuration of dots on the screen was the same as the one that had immediately preceded it in the 1-back condition, or as the one that had appeared two images previously in the 2-back condition (see Figures 2 and 3 for an example of each condition).

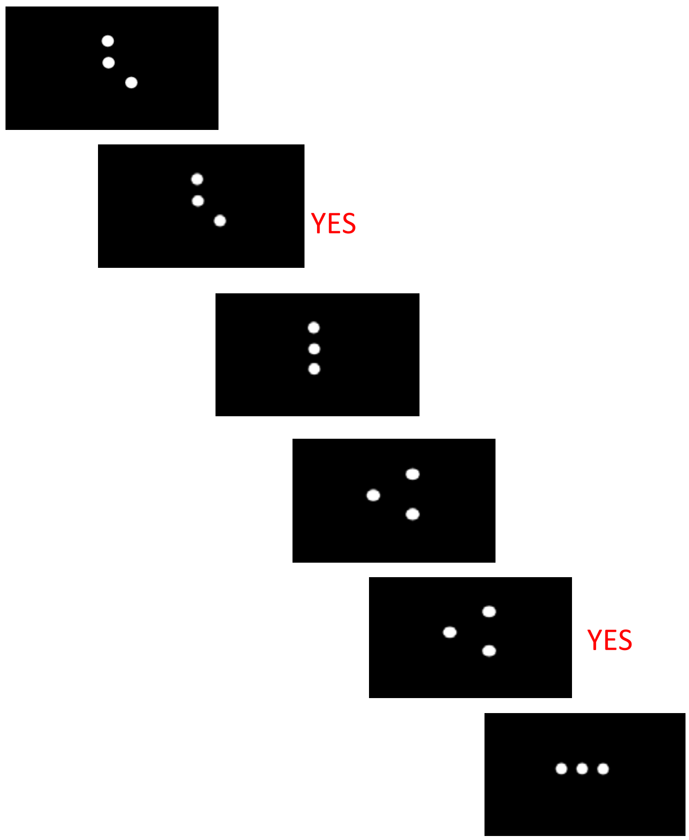

Figure 2: Partial example of a 1-back trial

In order to perform the task in the 1-back condition, the child had to hold one visual representation in memory (the target) and compare it to the subsequent image, replacing each target image with the newest and now relevant image as the task progressed. In the more difficult 2-back condition, the child was required to hold two different visual images in memory (the target and the next image), compare the currently displayed image with the dot configuration presented 2 items previously, and then update the pair of images to keep in mind as each new image was presented.

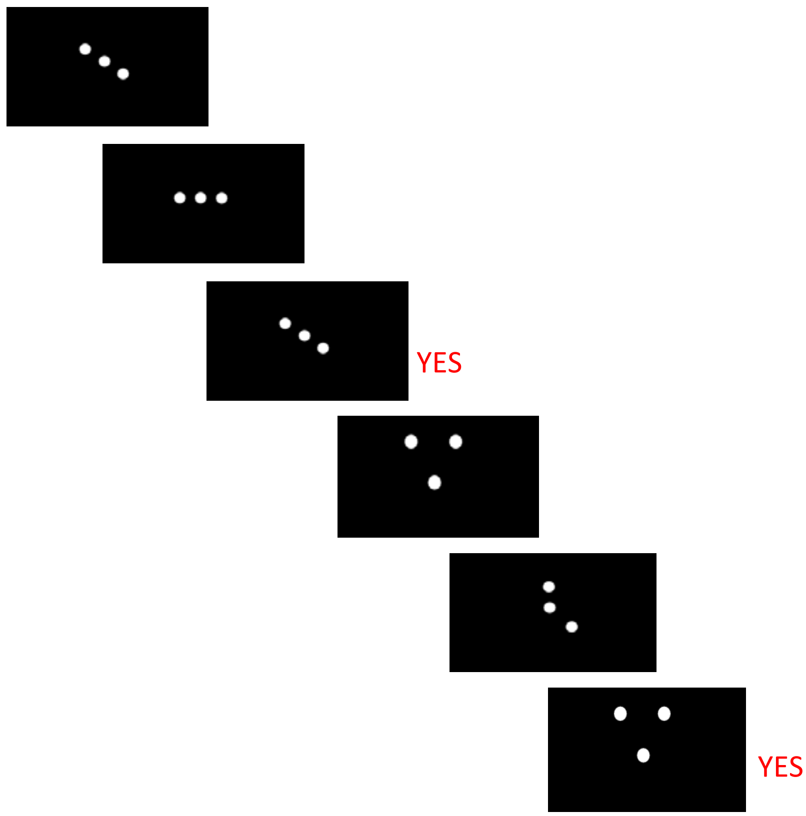

Figure 3: Partial example of a 2-back trial

Each participant completed three 1-back trials on the first day of testing, and three 2-back trials on the second day of testing. Each test trial consisted of 18 three-dot configurations, with six or seven hits (i.e., identical to the target) in each trial, for a total of 20 possible hits out of 54 responses in each of the 1-back and 2-back conditions. Prior to the test trials, a practice trial of six images (including two hits) was presented to the child on cue cards. This was followed by a demonstration trial of nine images (with three hits) presented on the computer, which the child then completed as a practice trial. This final practice trial was repeated once if the child did not succeed perfectly on the first attempt. If the child continued to make errors, the experimenter assisted the child and they repeated the practice together one last time. The experimenter recorded the child's responses online for later comparison with the electronic data recorded automatically by E-prime.

Given that the participants in the current study were younger than those from the few studies that have used the n-back task with children, and also based on the low success rates reported in Im-Bolter et al. (2006), the version developed here was simplified in a few ways. First, in each condition (1-back and 2-back), the total number of images (54 containing 20 hits) was split into three trials in order to reduce the attentional demands for the children and to allow the experimenter to repeat the instructions if necessary. Also, the children were required to respond by pressing the spacebar only when the image matched the target. This was deemed much simpler than having them associate one response key to 'yes' when the image was identical to the target and another response key to 'no' when it was not. Pilot testing confirmed that children in the youngest age group understood and were able to complete the 1-back task.

Initially, based on prior research, piloting, and the complexity of the 2-back condition, we expected that the 1-back condition would be sufficiently difficult and lead to variability in performance for children aged 5 to 8 years. However, as some of the children were doing very well, the 2-back condition was added to the experimental protocol in order to guard against any possible ceiling effects.

Signal detection theory "provides a general framework to describe and study decisions that are made in uncertain or ambiguous situations" (Wickens, 2002, p. 3). As such, it was the appropriate framework for reporting and judging performance on the n-back task. Each child's hit rate (h) and false alarm rate (f) were calculated separately for the 1-back and 2-back conditions. The hit rate is the ratio of 'yes' responses to the total number of times the stimulus was identical to the target (i.e., true positives). This corresponds to the proportion of accurate detections. The false alarm rate is the ratio of 'yes' responses to the total number of times the stimulus was ***not*** identical to the target (i.e., false positives).[3] Neither the hit rate nor the false alarm rate tells us the whole story. For instance, a respondent can be very conservative and answer only when he is absolutely sure that two images are identical. This would lead to a low hit rate, but also a low false alarm rate. On the other hand, another respondent may be more easily inclined to respond 'yes', which would lead to higher hit and false alarm rates. Finally, the best performance requires accurate detection, yet a degree of caution; this combination will result in a high hit rate ***and*** a low false alarm rate. By taking both the hit rate and the false alarm rate into account, a single value can be derived to represents each child's success on the n-back task. The most interpretable and simplest way to achieve this is to measure the difference between the hit rate and false alarm rate ($h - f$; Lockhart, 2000).

The *sound monitoring* task was adapted from the tone monitoring task developed by Miyake and colleagues (2000) for use with young children. It involves auditory presentation and a nonverbal response, and does not seem amenable to verbal mediation, at least by young children. The task was comprised of three sounds representing a bird, a bell, and a car horn. The experimenter was careful not to label the sounds in the instructions.

For each trial, the child listened to a sequence of nine sounds, three instances of each the bird, the bell, and the car horn arranged in random order and presented over headphones. The child responded by pressing the spacebar when he detected the third presentation of each sound. The sounds were 500 ms in duration, followed by 1500 ms of silence. The trials were programmed and presented using E-prime software, and each child heard the same trials in the same order.

---

[3] The response to each item can correspond either to a *hit* ('yes' when identical), a *miss* (no response when identical), a *false alarm* ('yes' when ***not*** identical) or a *correct rejection* (no response when **not** identical). However, given that *Age and scores for each participant on the memory tasks and for referential adequacy by story* i) the hit rate (h) and the miss rate ($1 - h$) are complementary, and that ii) the false alarm rate (f) and the correct rejection rate ($1 - f$) are complementary, the combination of the hit rate and the false alarm rate is sufficient to describe a child's pattern of responses.

Prior to the test trials, the participant observed as the experimenter completed a demonstration trial that was identical in structure to the experimental trials. The child then completed a practice trial, which was repeated once if it was not completed accurately. If the child continued to make errors, the experimenter and the child repeated the practice together one last time. The experimenter recorded the child's responses online, which were later compared with the electronic data recorded automatically by E-prime. The same analysis strategy was used as for the n-back task, as the number of correct responses (hits) and number of incorrect responses (false alarms) were recorded for each child, and then converted to a hit rate and a false alarm rate.

A modified *keep track* task (Yntema, 1963; Miyake et al., 2000) was used in this study as the final measure of children's updating abilities. This task involves visual presentation of nameable pictures and a verbal response, and thus strongly invites verbal mediation. The child was shown a series of pictures, each of which belongs in one of six categories (i.e., body parts, toys, food, clothing, animals, and transportation). The objective of the task was to remember the last picture presented in each of the two (or three) target categories. Prior to the picture presentations, the participant was told which two (or three) were target categories and icons representing these categories remained on the bottom of the screen throughout each trial. The pictures were then presented one at a time in the middle of the computer screen. Each picture was presented for 2000 ms, followed by 1000 ms of blank screen. A response screen appeared at the end of the trial with a question mark and the child was asked to recall the final item belonging in each of the target categories. This task required the child to update the contents of working memory each time a new item that belonged in one of the target categories was presented.

The stimuli included a total of 48 colour pictures, eight from each of the six categories.

The majority of the pictures were taken from the colour pictures developed by Rossion and Pourtois (2004) based on the Snodgrass and Vanderwart (1980) object set, and available as freeware at www.nefy.ucl.ac.be/facecatlab/stimuli.htm. A few additional pictures were taken from Boardmaker (Mayer-Johnson, Inc.). The category icons were developed from clipart line drawings. The objects selected all had one- or two-syllable labels, and were early-acquired words.[4]

A total of nine trials were completed, with three trials presented in each of three conditions of varying difficulty. The trials were constructed in order to insure that they were well-balanced in terms of how often each category and each picture appeared (as targets and foils), and syllable length. Given these constraints, the order or the pictures (targets and foils) was determined quasi-randomly. Each child received the same trials in the same order.

In condition 1, each of the three trials had two target categories and 10 pictures were presented. Of those 10 pictures, six were from the two target categories (three from each) and four foils were from the non-target categories. Condition 2 also included three trials and two target categories in each trial, but in this condition 12 pictures in total were presented, with four items from each of the two target categories, and three foils. This condition required the child to make one additional update for each target category being remembered. The final condition was the most difficult, as it added an additional target category. Each trial consisted of three target categories, and a total of 12 pictures, three from each of the target categories and three foils.

Foils were included in all trials to reduce the likelihood that participants would attend only to the pictures appearing towards the end of the presentation. Prior to the task, each child was shown the pictograms for all six categories, and the eight pictures belonging to each

[4] According to the Bates-MacArthur Communicative Development Inventories (CDI) American English norms (Dale & Fenson, 1996), 75% of children spoke the word prior to the age of 30 months; the norms are available online at http://www.sci.sdsu.edu/cdi/lexical. Objective age of acquisition data from a picture-naming study by Morrison, Chappell, and Ellis (1997) were used as complementary evidence when the words did not appear in the inventories.

category. The participant was asked to label each item and to select which category it belonged to. This pretest insured that the child was familiar with each picture and the category that it corresponded to. If the child labelled a picture incorrectly or associated a picture to the wrong category, the experimenter provided the correct response. This occurred very rarely. After the pretest, the experimenter completed a demonstration trial of eight pictures and two target categories, following which the child completed an equivalent practice trial. If the child did not succeed on the practice trial, it was redone once. If the child continued to make errors, the experimenter assisted the child and they repeated the practice together one last time. The experimenter recorded the child's responses online. An audio recording was also made to insure scoring accuracy. A single point was given for each item correctly recalled.

To elicit narrative productions, the wordless picture books *Frog on His Own* by Mercer Mayer (1973) and *April Fools* by Fernando Krahn (1974) were used. These two stories were chosen because they are comparable in that they are depicted in black and white pictures, both contain mainly human characters, and the two stories are comparable in length and complexity. *Frog on His Own* is about a pet frog that leaves his owner in a park and embarks on a journey of wreaking havoc for a number of other park users. *April Fools* was selected as a compliment to *Frog on His Own* because the story is centered on two boys as the main protagonists who play a prank on the townspeople by scaring them with a monster that they constructed. As this story has many groups of characters (i.e., two boys, townspeople), it provides an opportunity to examine how children move from referring to the group collectively, to singling out individual characters. It also differs in terms of having humans rather than an animal as the main protagonists. For each story, the child was given the book and instructed to flip through the pages to become familiar with the story. When the child was ready, the experimenter explained that the story would be recorded for a friend who was going to listen to the story later. The experimenter reminded the child that the listener would not be able to see the pictures. The experimenter then positioned herself a few feet away, where she could not see the pictures. The environment was set up this way in order to reduce the shared contextual information between the participant and the listener, and thus increase the likelihood that the child would be more precise when referring to the story characters. In the few cases where the participants had difficulty starting their stories, the prompt "tell me what's happening in the first picture" was used. The narratives were video-recorded for later transcription and coding.

The experimental procedures were divided into two separate testing sessions, each lasting approximately 30 minutes, and generally separated by one week. Day 1 of testing began with visual span, followed by the 1-back condition from n-back, and finally sound monitoring. Day 2 began with keep track followed by *April Fools,* digit span, and finally *Frog on His Own*. The 2-back condition was completed either at the end of day 2 or on a separate day for a few of the participants. All participants completed the tasks in the same order to avoid any possible order effects that might have been difficult to interpret, and given the relatively small sample size. There were no obvious order effects anticipated. The stories were given on the second day, when the children would be more comfortable with the experimenter.

2.2.2. Transcription and coding of the narratives

All stories were orthographically transcribed using the Systematic Analysis of Language Transcripts software (SALT; Miller & Inglesias, 2006). The stories were segmented into communication units (C-units) following the criteria outlined by Loban (1976). C-units consist of main clauses along with any dependent phrases and clauses. Coordinated clauses (using *and*, *but*, or) are treated as separate C-units except in cases where the co-referential subject of the second clause is omitted. Direct quotations consisting of more than one clause were segmented according to the same rule. Story closings (e.g., “the end,” “that’s it”), tangential comments, unintelligible or abandoned utterances, and mazes (including filler words, false starts, retraces, and repetitions) were excluded from the main story body and from the word count. A second listener reviewed all transcripts and also double-checked for utterance segmentation. Any disagreements were resolved by discussion.

Descriptive indices including mean length of C-units in words (MLCU-w), total number of C-units, and number of different words were calculated automatically using the Systematic Analysis of Language Transcripts program (SALT; Miller & Iglesias, 2006). Once the stories were transcribed and segmented, each mention of a story character was identified and coded on three dimensions: referential function, referential type, and referential adequacy.

The procedure used for classifying referential function was adapted by Wong and Johnston (2004) from the work of Bamberg (1987). Each mention of a story character was classified into one of the following three categories:

A. Introduction: A reference to a story character serves the introduction function when it is the speaker’s first mention of a story character.
B. Maintenance: A reference to a story character serves the maintenance function when the speaker continues to refer to the same character within an utterance or in successive utterances (in this case, C-units). An utterance that does not advance the thematic progress of the story, such as a descriptive statement, is not considered intervening; consequently, reference to characters can be maintained across such utterances.
C. Reintroduction: A reference to a story character serves the reintroduction function when the speaker refers back to a character who was previously introduced after intervening utterances focused on another character or on another aspect of the story.

Referential type corresponds to the lexical forms used to refer to story characters. This study used a classification system modified from the work of Halliday and Hasan (1976), and included the following four categories:

A. Personals: Personal pronouns (e.g., *he, she, they, it*).

B. Possessives: Possessive pronouns and determiners (e.g., *his, hers, their*).

C. Nominals: Proper names, as well as nouns accompanied by indefinite articles, definite articles, and demonstratives (e.g., *Bob, a boy, the dog, this frog*).

D. Ellipsis: Omission of a noun phrase.[5]

An additional layer of coding was completed involving referential type. Data was collected regarding the type of forms used to refer to main characters and secondary characters. In the *April Fools* story, the two boys and the monster were considered to be main characters, and all other characters were coded as secondary characters. For *Frog on His Own*, all characters were considered secondary characters with the exception of the frog, who was classified as a main character.

Finally, each reference to a story character was judged for referential adequacy, which refers to whether the child used a linguistically appropriate devise that resulted in clear understanding by the listener regarding who was being referred to. The classification system, based on work by Beliavsky (2003), Halliday and Hasan (1976), and Liles (1985a), identified each character reference as *complete, incomplete*, *ambiguous, or exophoric.* Complete references correspond to adequate use of the linguistic form and are considered cohesive. Incomplete, ambiguous, and exophoric references correspond to inaccurate use of the lexical form or to inappropriate ellipsis. Consequently, the character's identity is either unrecoverable from the

[5] Ellipsis was considered an instance of character reference only when the linguistic forms omitted were limited to the noun phrase. In cases when the ellipsis was more extensive, it became difficult to judge exactly what was part of the ellipsis even though it may have been grammatically and pragmatically appropriate.

text, or confusion exists as to which character the child is referring to. Each reference to a character was classified into one of the following categories:

A. Adequate Forms:

- Complete reference: Complete reference indicates that the listener can ascertain the character's identity easily and without ambiguity. Complete reference can occur at the moment of introduction or when a character is referred to again. In the case of maintenance or reintroduction, the linking device can refer to a character that was introduced either before (anaphoric) or after (cataphoric) the current mention. This type of reference is considered cohesive, because the listener is able to successfully determine the identity of the referent.

B. Inadequate Forms:

- Incomplete reference: Reference to a story character is judged to be incomplete if the information referred to by the referential device is not provided in the text. Generally, an incomplete tie occurs when a character is introduced using either a definite article or a pronoun, or when a character is reintroduced with a pronoun.
- Ambiguous reference: Reference to a story character is deemed ambiguous when the referential device may correspond to two or more characters mentioned elsewhere in the text, thereby making the identity of the character ambiguous for the listener. Reference classified as ambiguous includes cases where the pronoun used can refer to more than one character, reintroduction using an indefinite article, errors in lexical choice (e.g., switching from *he* to *it*; changing the name of a character), and inappropriate ellipsis.
- Exophoric reference: The referential device is exophoric if the speaker relies exclusively on the context of the situation and information outside the text in order to specify the identity of a character. For example, a child may point to the book and say "this boy". This type of reference is not considered complete because the listener cannot recover the identity of the

referent from elsewhere in the text.

2.2.3. Reliability

The stories of four children of the 37 (10%) were independently transcribed and segmented into C-units by a native English speaker. Word-for-word agreement for the transcripts was 92%. Agreement on utterance segmentation was 95%. Given the complexity of the scoring scheme and the large numbers of tokens, a second coder double-checked all transcripts for the coding of referential function and adequacy. All disagreements were discussed between the two coders, who came to a final decision. Reliability of scoring was estimated based on the number of coding changes in the samples of four children. Disagreements were rare, and the overall reliability for function and adequacy combined was estimated at 94%.

Accuracy of online scoring for the n-back and sound monitoring tasks was checked against the E-prime electronic scoring record; in cases where the computer program recorded a response that the experimenter did not, the computer record was taken as correct. Video or audio recordings were used to assess the reliability of the online scoring for the keep track, visual span and, digit span tasks. Based on 10% of the sample, reliability was found to be 100%. Given this high level of accuracy, the online scoring was assumed to be exact.

3. RESULTS

This chapter opens with a description of the characteristics of the narrative productions and a summary of the overall performance on the memory tasks. The results addressing each of the research questions will then be presented in turn, beginning with the primary analyses evaluating the relationship between updating and reference, followed by the results from the developmental analyses of updating and reference, and concluding with the analysis of the relationship between referential function and referential adequacy.

3.1. Story Characteristics

All participants told two stories from wordless picture books, *April Fools* followed by *Frog on His Own*. Descriptive data about story length, utterance length, and lexical diversity are presented in Table 2. The participants generally produced longer narratives with greater lexical diversity for *Frog on His Own*. On the other hand, mean length of utterance was more comparable for the two stories, but slightly longer for *April Fools*.

Table 2: *Means and ranges for number of utterances (measured in C-units), mean length of utterance in words, and number of different words, by story and by grade*

Story		Grade		
		K	1	2
April Fools	C-units	24.6 (7-40)	26.9 (9-50)	30.2 (15-83)
	MLCU-w	6.8 (2.9-8.4)	7.6 (5.1-11.4)	8.5 (7.4-9.8)
	NDW	66.8 (24-107)	82.5 (43-148)	88.5 (44-112)
Frog on His Own	C-units	32.2 (20-58)	32.8 (16-44)	40.6 (14-140)
	MLCU-w	6.3 (3.7-7.6)	7.2 (4.6-9.2)	8.3 (6.4-10.3)
	NDW	81.7 (37-122)	86.2 (62-111)	115.2 (49-343)

The participants in all age groups produced stories of considerable length, with mean story length increasing slightly from kindergarten to grade 2 (*April Fools*, 24.6 to 30.2 C-units; *Frog on His Own*, 32.2 to 40.6 C-units). They also made many references to story characters, thus providing many tokens to study referential adequacy (see Table 3). The fact that the participants made more references to characters in *Frog on His Own* was most likely due to the greater number of characters in this story compared to *April Fools*.

Table 3: *Means and ranges for number of references to story characters, by story and by grade*

Story	Grade		
	K	1	2
April Fools	42 (15-58)	47 (21-85)	63 (31-182)
Frog on His Own	52 (25-87)	60 (28-91)	81 (22-288)

To account for the differences in story length across children, referential adequacy was based on the proportion of complete references to story characters instead of on absolute numbers. Table 4 presents the mean referential adequacy data by grade and for each story. Referential adequacy was highly correlated across the two stories, r (37) = .70, $p < .01$, one-tailed.[6] As a result, the data for complete and total references to characters from the two stories were combined to create an overall referential adequacy score for each participant. The overall proportion of complete references was then transformed using the arcsine transformation before performing any statistical analyses. The arcsine transformation is recommended for inferential statistical analyses because the means and variances of proportional data are correlated, which makes them less suitable for analysis (Winer, 1991). It also serves to create greater dispersion in the tails of the distributions of scores (Judd, McClelland, & Ryan, 2009). In the specific case of the referential adequacy scores, the arcsine transformation resulted in a distribution of scores that better approximated a normal distribution and in a linear relationship between the adequacy scores and age.

Table 4: *Means (and standard deviations) for proportion of complete references, by grade*

Story	Grade		
	K	1	2
April Fools	.60 (.15)	.71 (.12)	.77 (.11)
Frog on His Own	.71 (.12)	.82 (.10)	.88 (.06)
Overall	.67 (.11)	.78 (.08)	.83 (.08)

3.2. Performance on the Memory Tasks

Results for the three updating tasks and the two short-term memory (STM) tasks are presented in Table 5, by grade. All the children in this study completed and were generally

[6] Unless it is otherwise specified, the significance levels reported for correlations are one-tailed, as there was an expected direction for the relationships tested. Given the relatively small sample size, this reduces the likelihood of committing a type II error.

successful on the keep track and sound monitoring tasks as well as the visual span and digit span tasks. All but two of the participants clearly understood and did generally well on the 1-back task. However, two first graders stood out: one who almost never responded (#66, 1 hit, 4 false alarms), and another who responded on almost every item (#64, 20 hits, 24 false alarms). Only three children (one in grade 1, and two in grade 2) were at ceiling, with 20 hits and no false alarms.

Table 5: *Means (and standard deviations) for performance on the updating and STM tasks, by grade*

Memory task	Grade		
	K	1	2
Keep track	11.3 (3.1)	12.0 (3.3)	15.8 (2.9)
1-back (*h-f*)	.80 (.11)	.82 (.17)	.90 (.07)
2-back (*h-f*)	.33 (.19)	.36 (.17)	.45 (.14)
Sound monitoring (*h-f*)	.53 (.39)	.79 (.17)	.86 (.13)
Visual span	5.2 (2.2)	7.9 (3.0)	8.2 (2.5)
Digit span	7.1 (1.5)	7.5 (1.3)	7.9 (1.6)

Note. Keep track, total number of items correct. 1-back, 2-back, and sound monitoring, hit rate minus false alarm rate. Visual span, mean span level. Digit span, total number of trials correct. For all but one variable, n = 12 for kindergarten, n = 13 for grade 1, and n = 12 for grade 2. For 1-back only, n = 11 for grade 1.

As expected, the children had much more difficulty with the 2-back condition, and many of them did quite poorly. Whereas the hit rates decreased overall, the false alarm rates generally increased considerably. The subjective impression of the experimenter during data collection was that many children were responding only on every second item and that others forgot what they were supposed to do as the trials went on. Additionally, scores on the1-back and the 2-back conditions were not significantly correlated ($r = .266$, $p > .05$).[7] These results suggest that the 2-back condition was too difficult for the majority of the children in this study. The 7- to 10-year-old children in a study by Im-Bolter et al. (2006) experienced similar difficulties with the 2-back

[7] This result is partially attributable to the fact that the two children who did so poorly in the 1-back condition, actually did much better in the 2-back condition.

condition, which resulted in the authors using only the data from the 1-back for statistical analyses. The same decision was made here, and only the data from the 1-back condition was used in the analyses. The data from the two participants whose performance clearly stood out was excluded from the analyses as they were not deemed representative of their abilities; this decision was also motivated by the fact that they were consistently identified as statistical bivariate outliers or as presenting high leverage values. Hence, unless it is otherwise indicated, all analyses involving the data from the 1-back task include the scores for 35 participants.

For each participant, the sound monitoring and the n-back tasks were scored for both hits (i.e., correct responses) and false alarms (i.e., incorrect responses), and these were converted into a hit rate (h) and a false alarm rate (f). Following the method suggested by Lockhart (2000), each participant's false alarm rate was then subtracted from his/her hit rate ($h - f$). This variable was then adjusted using the arcsine transformation. The final statistics entered into any analyses for the sound monitoring and the 1-back tasks were the arcsine of $h - f$ for each participant. Once again, these transformations resulted in a better distribution of scores, and linear relationships between variables. The data collected from the keep track task was the total number of items correctly recalled; no transformations were required for this data, and raw scores were entered into statistical analyses.

The updating measures were chosen to contrast performance between tasks that were deemed either visual, auditory, or verbal given the nature of the stimuli, the presentation and response modalities, and the extent to which they invited verbal rehearsal. All correlational analyses were preceded by visual inspection of the data. In addition, precautions were taken to identify bivariate outliers (using standardized residuals) and high leverage values.[8] The

[8] Any data point with a standardized residual value greater than |2.58| was considered an outlier, which corresponds to a 99% confidence interval and is a generally accepted cut-off point for small samples (Field, 2005; Tabachnick & Fidell, 2001). Any points with leverage values above .20 were considered potentially problematic (Zumbo, 2002). When necessary, follow-up sensitivity analyses were performed.

correlations between updating measures were all significant using adjusted *p*-values for the three comparisons based on the Holm method[9]: keep track and sound monitoring, r (37) = .355, $p <$ .017; sound monitoring and 1-back, r (35) = .350[10], $p < .025$; keep track and 1-back, r (35) = .302, $p < .05$. Comparing the children's performance on the different updating tasks provides interesting information regarding the domain-general or domain-specific nature of the updating process. The significant correlations between all updating tasks indicate that the tasks may have tapped a common underlying updating ability. The magnitude of the correlations between the various updating tasks in this study are consistent with the values found in previous research with both adults and children, where correlation coefficients between updating measures ranged from .15 to .43 (Miyake et al., 2000; St. Clair-Thompson, & Gathercole, 2006; van der Sluis et al., 2007).

3.3. The Relationship Between Referential Adequacy and Updating Skills

The relationship between a child's updating skills and his/her ability to adequately refer to story characters was evaluated using Pearson bivariate correlations. All three updating tasks were significantly correlated with referential adequacy using the Holm method to control for multiple comparisons.

Sound monitoring and referential adequacy	r (37) = .445, $p < .017$
1-back and referential adequacy	r (35) = .347, $p < .025$
Keep track and referential adequacy	r (37) = .333, $p < .05$

[9] Throughout the study, the Holm correction was chosen to control for Type I error when appropriate. This method is more powerful yet never rejects fewer comparisons than the Bonferonni procedure (Aickin & Gensler, 1996). As such, it proved to be a better choice to also guard against Type II error given the relatively small sample size. First, the *p*-values for all the comparisons are placed in increasing order. Then, each *p*-value is compared with $\alpha/(Nc - i + 1)$ for rejection of the null hypothesis, where Nc corresponds to the number of comparisons, and *i* to the rank of each specific comparison. No further tests are done beyond the first non-rejection.

[10] The results were essentially identical when the correlation between 1-back and sound monitoring was performed with one bivariate outlier removed, $r(34) = .370$, $p < .025$.

The moderate positive correlations between adequate reference and updating indicate that children's updating abilities likely contribute to being able to clearly refer to story characters, but they also suggest that updating is one factor among many that contribute to referencing abilities.

Although a relationship between updating and referencing skills was found, one could argue that the relationship may not be driven by the dynamic manipulation of working memory contents, but instead by a more basic memory capacity. To address this possibility, two STM memory tasks, digit span and visual span, were included as a measure of more basic memory storage capacity. Both a verbal and a visual span measure were included to reflect that fact that Baddeley's influential memory model has postulated the fractionation of STM into a verbal and a visuospatial component (Baddeley, 1996b). Additionally, both verbal and visual STM measures were selected to parallel the visual and verbal updating measures. As presented above, all three updating measures (auditory-verbal and visual) were found to be related to referencing abilities. The relationship was not as clear for the STM measures. The digit span did not show a significant correlation with referential adequacy, r (37) = .132, p = .219, whereas the visual span measure was significantly correlated with referential adequacy, r (37) = .346, p = .018. The lack of relationship between the digit span and referential adequacy suggests that the significant relationships found between both the keep track task and referential adequacy and the sound monitoring task and referential adequacy are not likely driven by a more basic verbal STM capacity. For the alternative explanation to hold, a larger correlation between verbal STM and reference would be expected. The significant correlation between visual span and reference, in the absence a relationship between verbal span and reference, was not predicted. Two possible explanations could account for this relationship; they will be briefly introduced here and expanded upon in the discussion. The first explanation is that digit span is not a good measure of verbal STM capacity, and that the nature of the task could be masking the relationship between verbal STM and referential adequacy, and/or between verbal STM and updating. The other possibility is that visual-spatial STM is more closely connected to the central executive and, as such, draws upon resources that are common with those required in the updating tasks and to adequately refer to characters in a story. This could be a more general relationship or one that is specific to the tasks (visual span and 1-back) used in the present study. If one considers that the visual span task was akin to another visual updating task, then the relationship between visual span and referencing becomes more plausible.

If the visual STM task is more closely connected to the central executive, then one would expect significant correlations between it and the updating tasks. This was in fact the case, as indicated in post-hoc analyses by the significant Pearson correlations between visual span and all three updating tasks using the Holm correction:

Visual span and 1-back	r (35) = .470, $p < .016$[11]
Visual span and keep track	r (37) = .434, $p < .025$
Visual span and sound monitoring	r (37) = .373, $p < .05$

Not only were the correlations significant between the visual STM scores and those on the three updating tasks, but also the magnitudes of the correlations were slightly greater than those found amongst the updating tasks. In contrast, the analyses did not result in any significant correlations between digit span and any of the updating measures, with the largest correlation emerging for digit span and keep track, $r(37) = .233$, $p = .083$.

To address the question of which factors best predicted referential adequacy, the three updating measures and the visual span were entered into a multiple regression analysis. The results of this analysis indicated that sound monitoring was the best predictor of referential adequacy, $R^2 = .20$. The addition of any other variables into the model added minimal predictive

[11] The results were similar and the correlation between visual span and 1-back actually increased when the analysis was performed with two bivariate outliers removed, $r(33) =.565$, $p = .001$.

power, increasing the R^2, but never sufficiently to result in a statistically significant change in R^2. For instance, entering sound monitoring first, followed by 1-back, keep track, and visual span in a single step resulted in the highest yet not significant increase of R^2 to .27. Although sound monitoring resulted in the greatest R^2, the other two updating tasks and visual span all had similar predictive power when entered individually, with R^2 values ranging between .11 and .20.

3.4. Developmental Analysis of Updating

To evaluate the developmental trends in children's updating abilities, correlations between updating and age were examined. Significant correlations were found between all three updating measures and age, indicating that the capacity to update the contents of working memory continues to develop between 5 and 8 years:

Keep track and age	r (37) = .530, $p < .017$
Sound monitoring and age	r (37) = .418, $p < .025$
1-back and age	r (35) = .398, $p < .05$[12]

Further analysis of the children's performance on the keep track task provided interesting information about the relationship between task demands and the updating process. The keep track task was designed to manipulate both the number of items to be updated at one time and the total number of times the information was required to be updated in a single test trial. Condition 1 presented three items from each of the two target categories requiring the children to update each target category twice. In comparison, condition 2 included four items from each of the two target categories, which required that the children update each target category three times. As shown in Table 6, the average proportion of correctly recalled items decreased from

[12] The results were similar and the correlation between 1-back and age increased when the data from one participant who was identified as an outlier was removed from the analysis, $r(34) = .471$, $p = .002$.

condition 1 to condition 2 for all three age groups (kindergarten, .71 to .47; grade 1, .74 to .59; grade 2, .85 to .81), indicating that number of times that the participants had to update the contents of working memory was a factor affecting memory performance. However, this drop was much smaller for the eldest group in the current study, indicating that for the second graders the difference between having to update the contents of working memory one more time for each of two categories did not present much more of a challenge.

Table 6: *Mean proportions (and standard deviations) of correctly recalled items for the three conditions of the keep track task, by grade*

	Grade		
	K	1	2
Condition 1	.71 (.16)	.74 (.19)	.85 (.15)
Condition 2	.47 (.30)	.59 (.21)	.81 (.22)
Condition 3	.47 (.14)	.44 (.23)	.65 (.19)

Condition 3 added an additional target category, which meant that the children needed to update the items recalled twice in each category. Although conditions 2 and 3 both required that the contents of working memory be updated six times, an additional item had to be recalled in the last condition. For both grade 1 and grade 2 participants, performance decreased between conditions 2 and 3 (grade 1, .59 to .44; grade 2, .81 to .65), indicating that the total number of items to remember also affected updating performance. In this case, the kindergarteners stood out, as they found conditions 2 and 3 equally difficult.

3.5. Developmental Analysis of Referential Adequacy

Consistent with previous research, a significant correlation was found between referential adequacy and age, r (37) = .637, $p < .001$. This strong correlation indicates that, between the ages of 5 and 8, children become better at selecting linguistic devices that lead to clear identification of story characters. It is interesting to note that the same significant relationships were observed for all three functions: introduction, $r(37) = .546$; maintenance, $r(37) = .661$; and reintroduction, $r(37) = .507$; all $ps < .003$.

Further analysis of the types of referential devices used by each age group served to verify whether the younger children relied more heavily upon nominal forms for character reference than did the older children. In practice, by far the most common linking device used by children other than nominals was the pronoun (either personal or possessive), and so the three other types (personals, possessives, and ellipsis) were collapsed and will be simply referred to as pronominal forms. As shown in Table 7, the hypothesis was not supported, as the distribution of pronominal and nominal forms changed only very little with grade in *April Fools*, and the use of pronominal forms actually decreased with grade in *Frog on His Own*. The more influential factor in the choice between using nominal forms or other referential types to refer to characters appears to have been the specific story, as pronominal forms dominated in all grades for *April Fools* (representing between 75% and 78% of the references to characters depending on the grade), whereas *Frog on His Own* elicited substantially fewer pronominal forms (between 40% and 52% of the references to characters, depending on the grade).

Table 7: *Mean proportions (and standard deviations) of pronominal forms, by story and by grade*

Story	Grade		
	K	1	2
April Fools	.77 (.10)	.75 (.09)	.78 (.10)
Frog on His Own	.52 (.10)	.45 (.15)	.40 (.11)

The more significant developmental change in terms of referential type was the increase in the proportion of correct use, with accuracy increasing from kindergarten to grade 2 for both pronominal (*April Fools*, .58 to .73; *Frog on His Own*, .64 to .85) and nominal forms (*April Fools*, .66 to .92; *Frog on His Own*, .78 to .91). In addition, the children were more successful in using nominal forms compared to other types of referential devices for both stories and across all three age groups (see Table 8).

Table 8: *Mean proportions (and standard deviations) of correct nominal forms and pronominal forms, by story and by grade*

Story	Referential type	Grade		
		K	1	2
April Fools	Pronominals	.58 (.15)	.68 (.13)	.73 (.12)
	Nominals	.66 (.25)	.84 (.12)	.92 (.08)
Frog on His Own	Pronominals	.64 (.15)	.79 (.12)	.85 (.10)
	Nominals	.78 (.09)	.85 (.10)	.91(.06)

The distribution of pronominal versus nominal forms was also analyzed to determine if the participants showed any evidence of using a thematic subject strategy (Karmiloff-Smith, 1980). This strategy would result in pronouns being reserved for reference to primary characters, with references to all secondary characters relying on nominal forms. The data presented in Table 9 indicates that, overall, the participants showed no tendency to use pronouns exclusively for reference to primary characters. The only large differences in the distribution of pronominal forms between primary and secondary characters appeared for the kindergarten children (primary .80, secondary .59) and the first graders (primary .79, secondary .65) for the *April Fools* story, but this trend was not evident in the data from *Frog on His Own* for any grade. Also, for both stories and all grades, pronouns were frequently used for secondary characters, with the pronominal forms representing between 39% and 76% of the references to secondary characters. Finally, there were no consistent patterns with increasing grade in terms of the children's tendency to rely on pronouns to refer to secondary characters; once again, the specific story seems to have exerted the most influence in terms of whether children chose to rely on nominals or other referential types to refer to either primary or secondary characters.

Table 9: *Mean proportions (and standard deviations) of pronominal forms used to refer to primary and secondary characters, by story and by grade*

Story	Characters	Grade		
		K	1	2
April Fools	Primary	.80 (.12)	.79 (.17)	.78 (.12)
	Secondary	.59 (.29)	.65 (.24)	.76 (.08)
Frog on His Own	Primary	.54 (.14)	.45 (.19)	.41 (.13)
	Secondary	.50 (.10)	.43 (.13)	.39 (.13)

3.6. The Influence of Referential Function on Referential Adequacy

Overall, in terms of referential adequacy measured as the proportion of complete references, the participants in all grades had the greatest success when maintaining reference to story characters (*M* .80, *SD* .08) and had much more difficulty when either introducing (*M* .51, *SD* .21) or reintroducing story characters (*M* .49, *SD* .15). Whereas reintroduction consistently posed the most difficulty for all grades for *April Fools,* the pattern was less consistent for *Frog on His Own*, with the kindergarteners and second graders presenting with the lowest level of referential adequacy for introduction, whereas reintroduction proved most difficult for the first graders (see Table 10).

Table 10: *Referential adequacy measured in proportion of complete references (means and standard deviations), by grade and by story*

Story	Function	Grade		
		K	1	2
April Fools	Introduction	.47 (.19)	.65 (.23)	.68 (.18)
	Maintenance	.80 (.13)	.84 (.11)	.88 (.07)
	Reintroduction	.30 (.21)	.48 (.27)	.54 (.26)
Frog on His Own	Introduction	.53 (.26)	.76 (.16)	.78 (.22)
	Maintenance	.79 (.10)	.89 (.09)	.94 (.04)
	Reintroduction	.66 (.13)	.72 (.18)	.84 (.10)

When considering the processing capacity required by the different referential functions, it was hypothesized that the various functions would vary in terms of the demands they placed on the working memory system. It was expected that reintroduction would show the highest correlation with updating, followed by maintenance, and finally introduction. To address this question, an analysis of the relationship between updating and referential function was completed. In order to limit the number of comparisons, the sound monitoring task was chosen to represent updating ability, given that it had the strongest relationship with overall referential adequacy. The largest correlation occurred between sound monitoring and maintenance $r(37) = .532$, $p < .017$, followed by reintroduction, $r(37) = .358$, $p < .025$, and then introduction $r(37) = .230$, $p = .085$, with only the first two meeting significance criteria using adjusted *p*-values for three comparisons based on the Holm method. It is interesting to note that for all three updating tasks, the strongest correlations consistently emerged for maintenance; this may partly be attributable to the larger number of tokens for this function compared to the two others. The average number of tokens for each of the functions across the three grades was 16 for introduction, 67 for maintenance, and 32 for reintroduction.

3.7. Summary of Results

- As a group, the children's updating scores showed a positive and moderate correlation with scores of referential adequacy. The strength of the correlations between the updating tasks and referential adequacy ranged from r = .333 -.445.
- The participants showed an increase in performance on the updating tasks with age, as evidenced by the moderate correlations between the three updating tasks and age (r = .398 - .530).
- Referential adequacy scores were strongly correlated with age (r = .637), indicating that from kindergarten to grade 2, the children became more successful at clearly referring to story characters. The distribution of referential forms by type did not change greatly with age, indicating that the children did not rely upon more specific nominal forms at a younger age.
- Overall, the children in this study did not exhibit evidence of a thematic subject strategy, since they often used pronouns to refer to secondary characters.
- The participants were more successful at maintaining reference to story characters, and had significantly greater difficulty introducing and reintroducing characters. Although reference overall was significantly correlated with updating, further analysis by function revealed that maintenance was most strongly correlated with updating (i.e. sound monitoring), followed by reintroduction, and finally introduction, which was not significantly correlated with updating.

4. DISCUSSION

The primary question motivating this study was to evaluate the developmental relationship between children's capacity to update the contents of working memory and their ability to accurately refer to story characters in narrative discourse. Previous research has provided evidence for a relationship between updating and language comprehension, but no study has explored the role of working memory updating during language production tasks.

Two narratives were elicited from each participant, and coding focused on reference to story characters. To measure children's capacity to update the contents of working memory, three tasks were designed to assess updating in both the visual and auditory-verbal domains. The primary results from this study are discussed first, followed by a discussion of the results from analyses of the developmental trends in updating and reference.

4.1. The Relationship Between Updating and Reference

The results from this study showed a moderate correlation between updating and referential adequacy. While the correlations found between the various updating tasks and reference (*r*s .33-.45) were not large, the strength of the correlations is nonetheless interesting and relevant, given that clearly referring to characters and memory updating both involve many complex cognitive operations. In addition, many levels of prior knowledge (including genre-specific, structural, and linguistic) are of course entailed in story production. Due to the multi-determined nature of the tasks examined in this study, large correlation coefficients would not be expected. The validity of the relationship between updating and reference is further supported by the consistency of the significant positive correlations found between each of the three updating tasks used in this study and referential adequacy. Due to the limitations of correlational analyses, the possibility that an additional variable could be driving the relationship between updating and reference cannot be ruled out. Both updating and character reference are abilities that are concurrently developing in the age group examined in this study and, as such, it is possible that age is driving the correlation. To address this issue, scatter plots illustrating the participants' scores on referential adequacy and each of the updating tasks were examined. The scatter plots showed overlap across the grades as well as a large degree of variability within each grade. Also, there was evidence for the general positive relationship between updating and referential adequacy within each grade, although it was not so clear for the second graders. This suggests that both a developmental trend and an individual differences factor may be present. Further support could be provided by future research with greater subject numbers that by increasing statistical power would allow for within-grade analyses, and make it possible to partial-out the age effects. It may in fact be possible that the strength of the relationship between updating and referential adequacy is changing over the age-range studied here.

4.1.1. Updating and referential function

This study is founded on the assumption that establishing and maintaining reference to story characters is a task that requires working memory resources and that the memory contents must be updated as the narrative progresses. Furthermore, it is possible that one of the referential functions (namely introduction, maintenance, or reintroduction) may place greater demands on the memory system. A further breakdown of the relationship between updating and each of the referential functions revealed the strongest correlations between updating and maintenance (r = .53) and between updating and reintroduction (r = .36) respectively, and a nonsignificant small correlation between updating and introduction. The lack of a significant relationship between updating and introduction makes logical sense. It may be driven by the referents that are introduced near the outset of the narrative, when there are as yet few characters to manage, and when minimal updating of the narrative model would have been required thus far.

Reintroduction, one could argue, would place the greatest demands on the working memory system, as the narrator must evaluate the listener's current level of familiarity with the character in question, consider which other characters have been introduced since the relevant character's first mention, and then carefully select a referring expression that will result in redirecting the listener's attention back to the intended referent.

The strong relationship between maintenance and updating is more difficult to interpret. Maintaining reference to story characters is the function children were most successful with, which is understandable give that maintenance by definition is the continuous mention of the same character over successive utterances. As the mentions are continguous, one would expect that this would not be overly demanding on memory processes. However, there were many instances in which multiple characters of the same gender were mentioned in the same utterance, which required that the narrator successfully disambiguate them for the listener through the use of cohesive devices in subsequent utterances. Finding a cohesive device that provided enough distinction between same-gender characters was a task that proved quite challenging, especially for the youngest participants. Nonetheless, the disproportionately larger numbers of instances where references to story characters corresponded to maintenance (on average 58% of the tokens) may in part be responsible for these findings.

4.2. The Development of Referencing Skills

The age range selected for this study was based on previous research that identified the early school years as the important window for the development of linguistic devices used for character reference. Referential adequacy scores from this study showed neither floor nor ceiling effects, and a strong correlation (*r* = .64) was found between referential adequacy and age. Together, these facts suggest that this was the appropriate age range to study. The changes in referential adequacy scores across the three grades studied showed a greater increase in accuracy between kindergarten and grade 1 (10%) than between grades 1 and 2 (5%). The level of referential adequacy found in this study is consistent with previous research. The kindergarten students achieved a mean overall referential adequacy score of 67% correct, which is similar to the 63% correct for the kindergarteners studied by Beliavsky (2003). Additionally, grade 2 students in the present study obtained an overall referential adequacy score of 83% correct, which is nearly identical to the results of Finestack, Fey, & Catts (2006) of 84% accurate referential ties for the same grade.

4.2.1. Evaluating referential accuracy by function

When examining the relationship between referential adequacy and referential function, the results from this study showed that maintenance was the easiest function for children to appropriately refer to characters for all three age groups and for both stories (87% accurate). Reintroduction was generally most difficult (62% accurate), although for *Frog on His Own*, the accuracy scores for introduction were slightly lower than for reintroduction for the kindergarteners and the second graders. There is not a large body of research that has examined the reintroduction function. Bamberg (1987), studied children's use of referring expressions for the switching function, where switching included both the introduction and the reintroduction of characters. He found that pronouns were often used inaccurately when the children switched between referring to different story characters. A study of Cantonese-speaking 3- to 12-year-old children showed similar patterns of data as those from the current study, with the reintroduction function having the lowest adequacy scores across all age groups (Wong & Johnston, 2004).

Additionally, the proportion of correct reintroductions was .38 for 5-year-old children and .66 for seven-year-old children in the study by Wong and Johnston, which is comparable to the results reported here, where the accuracy rates were .49, .62 and .68 for the kindergarten, grade 1, and grade 2 children, respectively. Although reintroduction was generally the most difficult function for children to adequately refer to story characters, there was a large discrepancy between the average reintroduction adequacy rates across the two stories. Overall, the participants were more successful with reintroduction for *Frog on His Own* than for *April Fools*. The same trend occurred with introduction and, to a lesser degree, with maintenance.

In order to make sense of these differences, it is important to look at the structure of each story. *April Fools* contains many characters that were referred to collectively, such as the two boys as the protagonists and the group of townspeople. As groups of characters all take the pronoun *they*, it becomes more challenging for the speaker to use cohesive referential devices while maintaining clarity of character identity. In addition, referring to one character that is part of the collective increases the level of difficulty, as additional information must be given to discern identity. Independently referring to one of the two boys in *April Fools* was frequently a challenge for the children in this study, and resulted in many ambiguous references. Therefore, the structure of the story used to elicit the narratives may have had as much influence on referential adequacy as did the variables of age and referential function.

In terms of the introduction function, the accuracy rates most likely reflect the dual influences of children's ability to master the distinction between indefinite and definite determiners and their understanding of which forms are appropriate for introduction. Previous research has identified indefinite and definite pronoun use as an area of difficulty for young children in similar contexts. Some studies have reported that 5- to 8-year-old children make many errors when introducing story characters and concluded that the indefinite/definite distinction is not mastered until the age of 9 or 10 (Hickman, 1980; Warden, 1976). The results from the current study revealed a substantial increase in adequate introductions from kindergarten to grade 1, as proportion correct increased from .47 to .65 for *April Fools* and from .53 to .76 for *Frog on His Own*. The results also point to an earlier age of acquisition of the appropriate linguistic forms for introduction when compared to earlier studies. For instance, Warden (1976) found that 7-year-old children used appropriate referring expressions for introduction only 61% of the time, and Hickman (1980) reported an even lower success rate of 49% at the same age. In comparison, the grade 1 (*M* age 6.9 years) and the grade 2 students (*M* age 8.1 years) in the current study achieved similar adequacy scores of 73% and 74% respectively across the two stories. In addition to the overall adequacy rates, it is interesting to note the levelling off of the accuracy scores, as indicated by minimal change between grade 1 and grade 2. One possibility is that this apparent plateau is followed by a further gain at a later age. A second possibility is that certain definite forms are accepted by the listener as adequate means for introducing characters in the specific discourse context used to elicit the narratives, even though researchers consider these forms incorrect. This suggestion has been discussed in previous studies that have found that even adult narrators do not always use indefinite forms when introducing story characters (Bamberg, 1987; Warden, 1976; Wigglesworth, 1990).

4.2.2. Development of pronouns and the thematic subject strategy

A large developmental change was seen in adequate pronoun use from kindergarten to grade 1, with a slower rate but additional improvement observed between grade 1 and grade 2. Similar to the effect of story that was found when examining referential function, the proportion of correct pronoun use is quite a bit greater in *Frog on His Own* as compared to *April Fools*. Once again, it seems likely that the challenges of referring to collective groups of characters and then discriminating among members of the group led to greater difficulty with pronoun reference in *April Fools.*

One hypothesis put forward in this study was that younger children may rely more heavily upon lexical cohesion, more frequently using nominal as opposed to pronominal forms. Pronouns and ellipsis provide much less detail as to the character's identity and, as such, could be more difficult for young language learners to use, resulting in their opting for simpler nominative expressions. This hypothesis was not supported by the results for either story, since a developmental change in the distribution of referential types was not found. Instead a developmental change was seen in the accuracy of use for both nominal and pronominal forms. Therefore, young children attempt to use the more difficult cohesive devices as frequently as the older children, although their accuracy is much lower.

In a cross-sectional study of the development of referring expressions, data collected from children 4 to 9 years of age lead Karmiloff-Smith (1980) to propose a developmental model that included a stage of pronoun use that she called the thematic subject strategy. According to this proposal, during this developmental stage, which appears around the ages of 6 or 7 years, children reserve pronominalization exclusively for the thematic subject, or main protagonist(s) in the story, with secondary characters being maintained using definite determiner + noun forms. This strategy was thought to aid the child in organizing his/her narrative. Bamberg (1987) found some evidence for a thematic-subject strategy for the 3- to-4 year-old children, but not for the 5- to 6-year-old children in his study, which contradicts Karmiloff-Smith's proposal that the thematic subject strategy emerges around the age of 6 or 7. The data from the current study provided minimal support for of the presence of a development stage during which children make use of a thematic subject strategy. Data from the kindergarten and grade 1 students showed a greater use of pronominal forms for reference to primary characters when telling the *April*

Fools story, but this discrepancy was not seen for the grade 2 children in *April Fools* or for any of the age groups when telling *Frog on His Own*. Therefore, the data from this study does not suggest that character pre-eminence (main versus secondary) influenced the speaker's selection of a specific type of linguistic form to refer to characters. As mentioned by Wigglesworth (1990), the characteristics of the story, such as how active the secondary characters are, as well as the length of time they remain in the story, likely impact the use of one referential type rather than another much more than character pre-eminence. Finally, the fact that the main protagonist in *Frog on His Own* was an animal may also have played some part in these results.

4.3. The Nature of the Updating Process

The correlations observed among the updating tasks in this study ranged between .35 and .37, which is consistent with previous research (.15-.43; Miyake et al., 2000; St. Clair-Thompson, & Gathercole, 2006; van der Sluis et al., 2007). The moderate correlations found between the updating tasks suggest that they were potentially tapping into a common updating ability. One might expect that if the tasks were measuring the same underlying ability, then even larger correlations would result. Because updating is a control process that is responsible for the manipulation of information in memory, it suffers from the task impurity problem discussed by many researchers in the field. Other processes that may be involved in central executive memory tasks include non-executive processes such as processing speed, other executive functions such as inhibition and attention, and additional factors such as prior knowledge and strategy use. In the present study, steps were taken to minimize the influence of other executive and non-executive resources; these included placing the 1-back and the 2-back tasks on separate days to eliminate the need to inhibit the previous response criterion, and shortening task condition length to reduce the demands on attention. Even with these precautions, it is likely that several extraneous cognitive processes are implicated in the updating tasks and contribute to the difficulty in achieving correlations of higher magnitude between updating measures.

The three conditions of the keep track task were included to further our understanding of the updating process by measuring the effect of altering task difficulty across the three conditions. The number of items to be updated at once and the total number of time the information was required to be updated in each trial were varied across the three conditions in an attempt to manipulate the load on the central executive. Comparisons can be drawn between condition 1 and condition 2, where condition 1 required two instances of updating for each target category and condition 2 required three for each target category. The second manipulation allowed for comparisons between condition 1 and 3, as the number of target categories increased from two to three. Results showed that updating each item just one more time added enough additional load to the central executive to result in decreased performance between condition 1 and condition 2 across all three age groups. The second manipulation of task load, the increase from two to three target categories, also impacted participants' ability to update working memory. Comparisons of results from condition 3 and condition 1 showed a substantial decrease in the number of items correctly recalled. Additionally, the discrepancy between conditions 1 and 3 was greater than that between conditions 1 and 2, indicating the number of items to be updated at one time affected task difficulty more than the total number of times updating was required in a single trial. The large impact that the number of items to be updated has on task accuracy is supported by results from the n-back. The 1-back condition required the child to hold one image in mind at a time as opposed to the 2-back condition, in which two images had to be held in memory at one time. The number of times the information had to be updated in each trial was consistent across both conditions. The increase from having to remember a single item at a time to two items simultaneously (and to update only one of the items each time) increased the task difficulty so substantially that most of the children were no longer successful. Previous research has provided concurrent evidence for both children and adults, that the central executive is a limited capacity system, and as the load placed on the updating process increases, performance accuracy decreases (Carretti et al., 2005; Yue et al., 2008).

4.3.1. Evidence for a domain-general updating capacity

The results from this study were interpreted within the framework of the Baddeley working memory model, as this is the most influential model currently used in memory research. Updating is a process that is believed to be part of the domain-general central executive (Baddeley 1996a, 2003b). Therefore, the resources of the central executive are not specific to the type of material nor to the modalities of presentation and response. To investigate the domain generality of the updating process, three different updating tasks were selected for this study to tap into both auditory-verbal and visual domains. Keep track and sound monitoring are presumed to load more heavily in the auditory-verbal domain, with n-back tapping into visual resources. Near equivalent correlations were found among the three different tasks, which supports updating as a domain-general process. The keep track task correlated as strongly with 1-back as it did with sound monitoring, further supporting a domain-general updating capacity. Furthermore, the relationships found between updating and reference, specifically the significant correlation between the 1-back task and reference, where the degree of the correlation was stronger than the correlation between the clearly more verbal keep track task and reference, provides additional evidence to support a domain-general view of updating.

4.3.2. The relationship between STM and reference

The correlation found between visual span and referential adequacy was initially quite difficult to interpret, especially in light of the absence of a significant relationship between referential adequacy and verbal STM as measured by digit span. In contemplating these findings, three different possibilities were considered. First, further review of the literature uncovered evidence of a closer relationship between the visuospatial STM system (the visuospatial sketchpad) and the central executive, when compared to the link between the verbal STM system (the phonological loop) and the central executive (Miyake et al., 2001; Morris & Jones, 1990; Yue, et al., 2008). Miyake and colleagues have proposed that visuospatial STM tasks are more difficult for participants because they are less practiced, less automatic, and less amenable to memory strategies (e.g., verbal rehearsal) when compared to verbal STM tasks. As a result, participants presumably have to call upon central executive resources in order to complete the visuospatial STM tasks, making it more akin to an updating task. Also, finding tasks across visual and verbal domains that are comparable in their level of difficulty and that result in similar degrees of variability has proven challenging, and the present study is no exception.

A second possible explanation resides in the design of the specific STM and updating tasks for this study. Further scrutiny of the tasks included in this study revealed that the visual STM task is quite similar to the visual updating task in terms of both structure and demands. In the case of the visual span task, the first image presented had to be remembered in order to be compared to the response image. The requirements of the 1-back task are similar in that the child had to remember the first image displayed for comparison to the next image, with the difference between tasks being that the second image also had to be remembered in the case of 1-back as it became the new target. Therefore, regardless of whether the visuospatial sketchpad is more closely linked to the central executive, the tasks used in this study may have been similar enough to be measuring the same process, thereby resulting in correlations with referential adequacy of comparable magnitude. Further research using additional visual STM and updating tasks will be required to tease apart these two interpretations.

The final point regards the lack of relationship found between the measure of verbal STM (digit span) and referential adequacy. It is possible that verbal STM is not a critical factor for explaining children's accuracy with referring expressions. Daneman and Carpenter (1980) failed to find a relationship between scores on the digit span and reading comprehension abilities in adults. The authors proposed that storage capacity is not the critical factor, but instead that it is the ability to utilize one's resources that is crucial for success in many language tasks. Contrary to Daneman and Carpenter's view, several research studies have found verbal STM to be critical for certain language abilities such as vocabulary acquisition (e.g. Baddeley, 2003a). In conclusion, results from this study suggest that passive storage of verbal information, as measured by the digit span, is not a crucial component for explaining individual differences in referential adequacy.

4.4. Future Directions

Although limited research has examined the cognitive operations of the central executive, many researchers are in agreement that such a central control process does exist. Baddeley conceptualized the central executive as part of his working memory model, but other researchers have explained the concept of a central executive as an even more basic function of controlled attention, or capacity for controlled processing (e.g., Engle, Tuholski, Laughlin, & Conway, 1999). The concept of controlled attention is not exclusive to the memory system but implicates cognitive functions such as inhibition and sustained attention. The positive correlations found between working memory measures and many higher-level cognitive tasks has lead Engle and colleagues to question which aspect of working memory is responsible for the correlations. Their answer is that it is not storage or memory, but rather the capacity for controlled attention. Strong correlations have been found between the commonly used reading span measure of working memory and several measures of memory updating, including the n-back and a numerical version of the letter memory task (Lehto, 1996; Schmiedek, Hildebrandt, Lovden, Wilhelm, & Lindenberger, 2009). Given the strong correlations between the complex working memory tasks and the updating tasks it is possible that these two measures are strongly related because they are both tapping into the capacity of controlled attention. The proposal of a more general controlled attention capacity, as opposed to a memory-specific central executive, is supported by the ability of working memory measures to predict performance on sustained attention and inhibition tasks (Turner & Engle, 1989). Therefore, although Miyake and colleagues (2000) found evidence for diverse executive functions (i.e. shifting, updating and inhibition), others have proposed controlled attention as the critical factor underlying performance on many executive functioning tasks. Further research is needed to determine exactly what updating tasks are measuring, and given the complex nature of the updating tasks, it is necessary to determine which component is most important for success on higher-order cognitive tasks.

A strong and consistent finding from this study is the significant influence that story structure can have on the choice and distribution of linguistic devices used for reference. Many existing studies have had children produce narratives from a single wordless picture book, leaving open the possibility that the characteristics of the story was in fact responsible at least in part for some of the results. The time required to elicit a story from a school-aged child is around 5 minutes; therefore, future studies should consider including a second book for elicitation because it is not time consuming and would allow for a more thorough interpretation of the results.

Due to the limited subject numbers in the current study, analyses by grade were not completed. It is possible that children need a certain level of updating capacity to be successful with character reference during narrative discourse, but that once the child has reached this threshold additional updating capacity does not result in improved character reference. Future research to increase the subject numbers would allow for the relationship between updating and reference to be evaluated in each grade to see if the relationship holds across each developmental group. Based on the current data, it is suspected that stronger relationships would be found in the younger age groups.

Previous research has established the relationship between working memory and language comprehension, and recent studies have examined the relationship between language comprehension and updating. The current study is, however, the first to evaluate the importance of memory updating for a specific aspect (reference to characters) in an online language production task. It is likely that updating is important for other aspects of language production and further research is needed in this area. Additionally, it would be interesting to study the importance of memory updating for language competence through a comparison between the typically-achieving children who participated in this study and children with specific language impairment.

4.5. Conclusion

The results from this study found that the ability to update working memory was related to referential adequacy in children 5 to 8 years of age. Due to the limited sample size and the limits of correlational analyses, these results are preliminary and would need to be replicated. The expected developmental progression was found in both updating ability and referential adequacy scores. Introduction and reintroduction of characters were the most difficult functions for the participants to manage in their narrative productions. Further investigation is needed into the relationship between updating and the various referential functions, particularly given that the strongest relationship was found between maintenance and updating, which was unexpected. The elicitation context is an important factor that can influence the narrator's use of referring expressions, but as found in this study, the specific story used to elicit the narrative also had a strong effect on both the type and adequacy of referential forms, a factor that may have been underestimated in previous research.

BIBLIOGRAPHY

Aickin, M., & Gensler, H. (1996). Adjusting for multiple testing when reporting research results: The Bonferroni vs Holm methods. *American Journal of Public Health, 86*, 726-728.

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable? *Child Development, 77,* 1698-1716.

Baddeley, A.D. (1996a). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49A,* 5-28.

Baddeley, A. D. (1996b). The fractionation of working memory. *Proceedings of the National Academy of Science, 93*, 13468-13472.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 4*, 417-423.

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

Baddeley, A. D. (2003b). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*, 829–839.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), The psychology of learning and motivation (Vol. Vol. 8, pp. 47-90). San Diego, CA: Academic Press.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 28-61). New York: Cambridge University Press.

Bamberg, M. (1987). *The acquisition of narratives: Learning to use language.* Berlin: Mouton de Gruyter.

Beliavsky, N. (2003). The sequential acquisition of pronominal reference in narrative discourse. *WORD,* 54, 167-189.

Bennett-Kastor, T. (1983). Noun phrases and coherence in child narratives. *Journal of Child Language, 10,* 135-149.

Berman, R., & Slobin, D. I. (1994). *Relating Events in Narrative: A Crosslinguistic Developmental Study.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, R. (1973). *A first language, the early stages.* Cambridge: Cambridge University Press.

Brown, G., & Yule, G. (1983). *Discourse analysis.* Cambridge: Cambridge University Press.

Carretti, B., Cornoldi, C., De Beni, R., & Romano, M. (2005). Updating in working memory: A comparison of good and poor comprehenders. *Journal of Experimental Child Psychology, 91,* 45-66.

Cornish, F. (1999). *Anaphora, discourse and understanding: Evidence from English and French.* Oxford: Clarendon Press.

Dale, P.S. & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, 28, 125-127.

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

Diamond, A. (2006). The early development of executive functions. In E. Bialystok & F. Craik (Eds.), *Lifespan cognition: Mechanisms of change.* (pp.70-95). Oxford, UK: Oxford University Press.

Emslie, H. C., & Stevenson, R. J. (1980). Preschool children's use of the articles in definite and indefinite referring expressions. *Journal of Child Language, 8*, 313-328.

Engle, R. W., Kane, W. J., & Tuholski, S. W. (1999). In. A. Miyake & P. Shah (Eds.), *Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex.* (pp.102-134). Cambridge: Cambridge University Press.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology, 125*, 309-331.

Field, A. (2005). *Discovering statistics using SPSS (and sex, drugs and rock 'n' roll) (*2nd ed.). London: Sage publications.

Finestack, L. H., Fey, M. E., & Catts, H. W. (2006). Pronominal reference skills of second and fourth grade children with language impairment. *Journal of Communication Disorders, 39,* 232-248.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive statuses and the form of referring expressions in discourse. *Language, 69* (2)*,* 274-307.

Grice, P. (1975). Logic in conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics, Vol. 3: Speech Acts*. (pp. 41-58). New York: Academic Press.

Halliday, M.A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman

Hickman, M. (1980). Creating referents in discourse: A developmental analysis of linguistic cohesion. In J. Kreiman & A. Ojeda (Eds.), *Papers from the parasession on pronoun and anaphor.* Chicago, Chicago Linguistic Society.

Hitch, G. J. (2006). Working memory in children: A cognitive approach. In E. Bialystok & F. Craik (Eds.), *Lifespan cognition: Mechanisms of change.* (pp.70-95). Oxford, UK: Oxford University Press.

Im-Bolter, N. Johnson, J., & Pascual-Leone, J. (2006). Processing limitations in children with specific language impairment: The role of executive function. *Child Development,* 77, 1822-1841.

Johnson-Laird, P.N., & Garnham, A. (1980). Descriptions and discourse models. *Linguistics and Philosophy, 3,* 371-393.

Judd, C. M., McClelland, G. H., & Ryan, C. S. (2009). *Data Analysis: A Model Comparison Approach.* New York, NY: Routledge.

Karmiloff-Smith, A. (1985). Language and cognitive processes from a developmental perspective. *Language and Cognitive Processes, 1,* 61-85.

Karmiloff-Smith, A. (1980). Psychological processes underlying pronominalization and non-proniminalization in children`s connected discourse. In J. Kreiman and A. E. Ojeda (Eds), *Papers from the parasession on pronouns and anaphora.* Chicago: Chicago Linguistics Society.

Kessler, Y., & Meiran, N. (2008). Two dissociable updating processes in working memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 34 (6),* 1339-1348.

Krahn, F. (1974). *April fools.* New York: Dutton.

Lehto, J. (1996). Are executive function tests dependent on working memory capacity? *The Quarterly Journal of Experimental Psychology, 49A,* 29-50.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge: MIT Press.

Liles, B. Z. (1985a). Cohesion in the narratives of normal and language-disordered children. *Journal of Speech and Hearing Research, 28*, 123-133.

Liles, B. Z. (1985b). Production and comprehension of narrative discourse in normal and language disordered children. *Journal of Communication Disorders, 18*, 409-427.

Loban, W. (1976). Language development: K through Gr. 12. Research Report No. 18, Urbana, IL: National Council of Teachers of English.

Lockhart, R. S. (2000). Methods of memory research. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 33-57). Oxford: Oxford University Press.

Logie, R.H., & Pearson, D.G. (1997). The inner-eye and the inner scribe of visuo-spatial working memory: Evidence from developmental fractionation. *European Journal of Cognitive Psychology, 9* (3), 241-257.

Mantyla, T., Carelli, M. G., & Forman, H. (2007). Time monitoring and executive functioning in children and adults. *Journal of Experimental Child Psychology,* 96, 1-19.

Maratsos, M. P. (1974). Preschool children's use of definite and indefinite articles. *Child Development, 45 (2),* 446-455.

Mayer, M. (1969). *Frog where are you?* New York: Dial Press.

Mayer, M. (1973). *Frog on his own.* New York: Dial Press.

Miller, J. Inglesias, A. (2006). Systematic Analysis of Language Transcripts (SALT), English and Spanish (Version 9) [Computer Software], Language Analysis Lab, University of Wisconsin-Madison.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex frontal lobe tasks: A latent variable analysis. *Cognitive Psychology,* 41, 49-100.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology, 130,* 621-640.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81,* 111-121.

Morrison, C.M., Chappel, T.D., & Ellis, A.W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *The Quarterly Journal of Experimental Psychology, 50A* (3), 528-559.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object set: The role of surface detail in basic-level object recognition. *Perception, 33*, 217-236.

Roth, F. P., Spekman, N. J., & Fye, E.C. (1995). Reference cohesion in the oral narratives of students with learning disabilities and normally achieving students. *Learning Disability Quarterly, 18,* 25-40.

Schmiedek, F., Hildebrandt, A., Lovden, M., Wilhelm, O., & Lindenberger, U. (2009). Complex span versus updating tasks of working memory: The gap is not that deep. *Journal of Experimental Psychology: Learning, Memory and Cognition, 35 (4),* 1089-1096.

Schneider,W., Eschman, A., & Zuccolotto, A. (2002). E-Prime (Version 1.1) [Computer software]. Pittsburgh, PA: Psychological Software Tools.

Semel, E.,Wiig, E.H., & Secord,W. A. (2003). *Clinical evaluation of language fundamentals, fourth edition (CELF-4).* Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology*, 125, 4-27

Snodgrass, J.G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 174-215.

St. Claire-Thompson H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology, 59 (6),* 745-759.

Tabachnick, B. G., & Fidell, L. S. (2001). *Computer-assisted research design and analysis*. Boston: Allyn and Bacon.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependant? *Journal of Memory and Language, 28*, 127-154.

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading and arithmetic. *Intelligence, 35,* 427-449.

Warden, D. A. (1976). The influence of context on children's use of indentifying expressions and references. *British Journal of Psychology, 67 (1),* 102-112.

Wickens, T. D. (2002). *Elementary Signal Detection Theory*. Oxford: Oxford University Press.

Wigglesworth, G. (1990). Children's narrative acquisition: a study of some aspects of reference and anaphora. *First Language, 10*; 105-125.

Winer, B. J. (1991). *Statistical Principles in Experimental Design* (3rd ed.). New York: McGraw-Hill

Wong, M. Y. A., & Johnston, J. R. (2004). The development of discourse referencing in Cantonese-speaking children. *Journal of Child Language,* 31; 633-660.

Yntema, D. B. (1963). Keeping track of several things at once. *Human Factors,* 5, 7-17

Yue, Z., Zhang, M., & Zhou, X. (2008). Updating verbal and visospatial working memory: Are the processes parallel? *Chinese Science Bulletin, 53,* 2175-2185.

Zumbo, B. (2002). Linear models, applied regression, and related methods: A unified approach for statistical modeling of observational and experimental studies [Course notes, EPSE 596, correlational designs and analysis in educational research], The University of British Columbia.

APPENDICES

Appendix A: Raw data

Table 11: *Age and scores for each participant on the memory tasks and for referential adequacy by story*

Subject ID	Age	KT	1-back	2-back	SM	Visual span	Digit span	Reference AF	Reference Frog
51	71	13	.81	.63	.33	3.33	7	.63	.65
52	74	16	.71	.35	.11	5.33	10	.62	.63
53	72	9	.75	.06	.61	4.67	7	.33	.52
55	70	12	.77	-.01	.67	4.33	5	.55	.83
56	76	10	.92	.47	.94	5.33	8	.74	.83
58	68	13	.82	.40	1.00	9.33	7	.78	.79
59	72	15	.95	.19	1.00	9.33	7	.72	.82
510	72	12	.89	.41	.11	4.67	8	.33	.71
511	67	10	.82	.31	.00	3.67	8	.69	.87
512	66	8	.84	.21	.06	5.67	8	.58	.66
513	73	5	.54	.43	.72	1.67	5	.56	.58
514	67	13	.75	.50	.83	4.67	5	.71	.60
61	77	10	.94	.40	.39	7.33	7	.56	.80
62	85	14	.95	.51	.94	10.33	6	.67	.84
63	86	11	.95	.17	.78	7.33	6	.84	.92
64	79	9	.29	.29	.94	7.33	8	.55	.79
65	87	15	.39	.21	.72	6.33	8	.75	.96
66	83	17	-.07	.49	.78	11.33	7	.77	.87
67	84	16	.92	.50	.83	6.00	8	.65	.88
68	81	15	.77	.19	.89	9.33	9	.67	.83
69	86	14	1.00	.01	1.00	10.33	8	.81	.85
610	81	8	.69	.24	.78	7.33	5	.67	.77
611	80	11	.82	.56	.50	13.33	8	.54	.71
613	84	8	.77	.24	.83	3.33	7	.85	.87
614	78	8	.79	.12	.89	3.33	10	.91	.57
71	95	21	.82	.14	.94	6.33	7	.61	.73
72	96	16	.85	.34	.72	9.33	7	.73	.90
73	93	13	.97	.57	.83	8.67	6	.91	.93
74	101	14	.84	.51	.78	7.33	9	.61	.85
75	98	15	.83	.45	1.00	7.33	6	.84	.90
76	101	12	1.00	.47	.89	12.33	10	.78	.93
77	99	18	.91	.56	1.00	7.33	10	.85	.94
78	91	17	.95	.60	1.00	10.33	8	.66	.90
79	98	11	.86	.39	.61	3.33	8	.87	.89
710	95	17	.86	.33	.72	12.00	8	.89	.89
711	100	18	1.00	.61	.94	6.33	10	.73	.84
712	98	17	.94	.46	.83	7.67	6	.73	.90

Note. Age in months. Keep track (KT), total number of items correct. 1-back, 2-back, and sound monitoring (SM), hit rate minus false alarm rate. Visual span, mean span level. Digit span, total number of trials correct. Reference, proportion correct for *April Fools* (AF) and *Frog on His Own* (Frog).

Appendix B: Summary of correlations

Table 12: *Correlations between reference, updating, and memory span measures*

Measure	1	2	3	4	5	6
1. Reference	-	.333	.357	.445	.346	.132
2. Keep track		-	.349	.355	.434	.233
3. 1-back			-	.370	.565	.209
4. Sound monitoring				-	.373	-.012
5. Visual span					-	.133
6. Verbal span						-

Note. For all but one variable, *n* = 12 for kindergarten, *n* = 13 for grade 1, and *n* = 12 for grade 2. For 1-back only, *n* = 11 for grade 1.

Appendix C: Detailed scoring decisions for referential function and referential adequacy

Referential function

Introduction

- First mentions of a story character using a proper noun or an indefinite article + noun were coded as complete.
- Introductions using a pronoun or definite article + noun were coded as incomplete.

Maintenance

- Maintenance of a story character using a pronoun was coded as complete, unless the pronoun could refer to two or more possible characters, which made it ambiguous.
- The use of pronouns to maintain reference to story characters was generally considered adequate. It was nonetheless deemed necessary for the child to re-specify the identity of the character at regular intervals, otherwise the distance became too great between the cohesive tie and the pronoun, making it difficult for the listener to follow. Hence, the child was expected to further specify the identity of the character at each change in scene. In the case of successive pronoun use, the first pronoun in the scene was coded an incomplete unless the speaker had just specified the identity in the preceding utterance, or if he/she used cataphoric reference to further specify the identity later in the scene.

Reintroduction

- Reintroduction using a proper noun or a definite article + noun was coded as complete.
- If the narrator switched from referring to a group of characters collectively, to one of the characters individually, the individual character reference was coded as a reintroduction.
- Reintroductions using an indefinite article were coded as ambiguous.

- Reintroduction using a pronoun was coded as incomplete except in cases where the pronoun could be considered to refer to more than one possible character, in which case it was coded as ambiguous.

Additional considerations

- Any reference error (e.g., switching case or gender of pronoun) was coded as ambiguous.
- The use of vague terms (e.g. *somebody*, *everybody*, *people*) for introduction was coded as complete if the identity of the referent was further specified at another point in the same scene or if these terms were used in truly generic way (i.e., making a general statement). If not, the introduction was coded as ambiguous.
- If the speaker told the story in the first person, introduction or reintroduction using the first person singular pronoun was coded as complete.
- In *Frog on His Own*, the frog's owner and his dog were introduced in the first scene and did not return to the story until the final scene. In this case, the speaker needed to reintroduce these characters using more that a definite determiner + noun. The speaker needed to provide some additional information that indicated to the listener that the boy and the dog were the same characters as those initially introduced. If no additional information was provided the definite determiner + noun was coded as ambiguous.

Appendix D: Story details

Scenes

April Fools

1. The boys are making the dragon.
2. The boys scare a woman sleeping.
3. The townspeople see a monster on a house roof and the boys scare a man in his house.
4. The boys take the monster across the river and get lost in the forest. They are found by the townspeople.

Frog on His Own

1. The boy, his pets, and the frog arrive in the park; the frog stays behind.
2. The frog catches a bee.
3. The frog disrupts a couple's picnic.
4. The frog breaks a boy's boat.
5. The frog jumps in a baby's carriage, gets chased by a cat, and is saved by his owner and the dog.

Characters

April Fools

Main characters:

- Boy 1
- Boy 2
- Monster

Secondary characters:

- Townspeople
- Victim 1 (old lady)
- Victim 2 (man)

Frog on His Own

Main character:

- Frog

Secondary characters:

- Boy 1
- Dog
- Turtle
- Bee
- Fly
- Man
- Woman 1 (picnic)
- Boy 2 (boat)
- Mom 1 (boat)
- Baby
- Mom 2 (baby)
- Cat

Appendix E: Detailed task instructions

Visual Span

"We are going to start with an activity on the computer. In this game you will see squares. Some of the squares are going to be red. Then all the squares will disappear, when they come back one of the red squares will be gone. Your job is to point to the square that was red before. As we go there are going to be more and more squares. If you're not sure it is ok to guess. Just do your best. Ready?" (Complete practice. If child makes a mistake, model the correct response and have the child do the practice again). (After level one say) "Ok, there will be more squares."

1-back

"This is a computer game. You will see three dots on the screen. These dots can line up in different ways. Here is what they look like." (Show the 10 possible dot configurations on the computer).

"In this game your job is to watch for a picture of the dots that is the same as the picture that came just BEFORE it. Let's try some for practice." (Present card 1), "look, remember this" (flip over card 1, present card 2). "Here's the next one, is it the same as the one before? Right it's not the same. Now remember this one" (point to card 2 then flip it over, present card 3). "Is this one the same as the one before? Now remember this," (point to card 3, flip it over, present card 4) "is this one the same as the one before?" (Continue flipping over each card over asking the child yes, or no).

"Now let's do it on the computer. I'm going to do some first, you watch. I am going to press this button (point to space bar) if the picture of the dots is the same as the picture just BEFORE it". "Look, remember this. Is this one the same as the one before?" (Repeat, "look, remember…" instructions for first 3 images, then nod or shake head).

“Ok, now it`s your turn to do it on the computer. I want you to press this button (point to space bar) if the one you see is the same as the one that came before it. Just try your best. If you are not sure it’s ok to guess.” (Present practice set on computer. If the child does not get them all correct, repeat the practice again with him/her). “That was close, let’s do it one last time together”. (When practice trial is complete say) “good job! Now you’re going to do this game 3 times by yourself”. (Complete 3 trials. If a break is needed between trials repeat the instruction) “your job is to press this button (point to space bar) if the picture of dots on the screen is the same as the picture BEFORE it”.

(After trial 1 say) “great job! 2 more to go!” (After trial 2 say) “you are doing a great job! This is the last one.”

2-back

“This is a dot game kind of like the one we played before but the rules are going to be a little bit different. In this game your job is to watch for a picture of the dots that is the same as the picture that came two pictures BEFORE it. Let’s try some for practice.” (Present card 1), “look” (present card 1), “remember this” (flip over card 1, present card 2) “we have to remember this one too and we’re still remembering this one” (point to card 2). “Now look at this one” (flip over card 3), “is it the same as the one you saw 2 pictures ago? (point to card 1). “We are remembering this one” (card 2), “want to peek? Now we have to remember this one too” (pointing to card 3, then flip over card 4). “Is this one the same as the one before?” (Continue flipping over each card over asking the child yes, or no).

“Now let’s do it on the computer. I`m going to do some first, you watch. I am going to press this button” (point to space bar) “if the picture of the dots is the same as the picture that came 2 pictures BEFORE it. Look, remember this. Ok, now we have to remember this one too. Now is this picture the same as the picture that came 2 before?" (Repeat, "look, remember" instructions for first 3, then nod or shake head. Experimenter completes the practice).

"Ok, now it`s your turn to do it on the computer. I want you to press this button" (point to space bar) "if the one you see is the same as the one that came 2 pictures before it. Just try your best. If you are not sure it's ok to guess". (Present practice set on computer. If the child does not get them all correct, repeat the practice again with him/her). "That was close, let's do it one last time together".

(Once the practice is complete say) "good job! Now you're going to do this game 3 times by yourself". (Complete 3 trials. If a break is needed between trials repeat the instruction) "your job is to press this button" (point to space bar) "if the picture of dots on the screen is the same as the picture that came 2 pictures BEFORE it".

(After trial 1 say) "great job! 2 more to go". (After trial 2 say) "great job. This is the last one!"

Sound Monitoring

(Open Sound Monitoring in E-Prime, both experimenter and child wear headphones). "We are going to play a game with sounds. These are the three sounds". (Play the three sounds for the child. DO NOT NAME THEM). "You are going to hear these sounds lots of times, <u>wait</u> until you hear the <u>same</u> sound 3 times, then you press this button" (point to space bar). "I will do the first one to show you. Ready?" (Experimenter completes the demo as the child listens. The child then completes the practice. If the child makes any mistakes experimenter repeats the practice with the child). "That was close! Let's do it one last time together." (When practice is complete say) “good job. Now you will do it by yourself. Wait until you hear the same sound 3 times, then press the button”.

Keep track

(Open in Power Point). “We are going to start with a computer game. In this game there are going to be lots of different things that belong in 6 groups”. (Show each icon as you name the category). “Now I am going to show you all of the things that go in these groups. You tell me what it is and then click on the group that the picture belongs to. Ready?” (Ask “what’s this?” as necessary. If child makes an error, stop and redo that object).

“Now I am going to show you how to play this game. My two groups are animals and clothing. We are going to see some of the pictures and at the end I have to say the last animal that I saw and the last piece of clothing that I saw, so those are the ones I have to remember. I am going to do this one, you watch, and then it will be your turn”.

“Now it is your turn. Your two groups are body parts and toys. At the end you have to tell me the last body part that you saw and the last toy that you saw. Ready?” (If child makes any mistakes say) “that was so close, let’s try it again” (and redo practice once more together with the child). “Great job! Now we are going to do another one. This time your groups are animals and clothing. Ready?” (Go through the first 6 blocks. Provide the child with the categories before each block begins. At block 7 say,) “now it is going to get a little bit tricky. You have to remember 3 groups. This time your groups are transportation, animals and toys. At the end you tell me the last transportation picture that you saw, the last animal that you saw and the last toy that you saw. Ready?”

April Fools

"I'm going to show you a book that has only pictures, no words. I want you to look through the pictures and think of a story to tell that goes with the pictures. When you are ready you can start telling your story. We are recording your story for my friend Sarah. She will listen to it later but she won't have the book. You tell me when you are ready" (child looks through pictures). "Ok tell the best story you can and I'll record it for my friend Sarah."

Digit Span

"I'm going to say some numbers. Listen carefully, and when I'm finished, you say them in the same order as I said them. For example, if I say 1, 2, you say 1, 2." (Discontinue when the child responds incorrectly on both items at a single level).

Frog on His Own

"Now we are going to tell another story. This time we are going to use this book. It is another book that has no words, only pictures. I want you to look through the pictures and think of a story to tell that goes with the pictures. When you are ready you can start telling your story. Just like before we are recording this story for my friend Sarah. She will listen to it later but she won't have the book. You tell me when you are ready" (child looks through pictures). "Ok tell the best story you can and I'll record it for my friend Sarah".

Appendix F: Word lists for the keep track task

Toys:	**Transportation:**	**Clothing:**
puzzle	car	shoes
doll	boat	dress
blocks	bike	hat
kite	bus	belt
balloon	airplane	shirt
crayon	truck	socks
book	tractor	pants
ball	train	sweater

Food:	**Body Parts:**	**Animals:**
apple	arm	cat
bread	ear	dog
cake	eye	frog
carrot	finger	duck
corn	foot	horse
grapes	hand	turtle
orange	lips	snake
sandwich	nose	pig

Appendix G: UBC Research Ethics Board certificate of approval

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 - FULL BOARD

PRINCIPAL INVESTIGATOR:	INSTITUTION / DEPARTMENT:	UBC BREB NUMBER:
Judith R. Johnston	UBC/Medicine, Faculty of/Audiology & Speech Sciences	H09-00350

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:

Institution	Site
UBC	Vancouver (excludes UBC Hospital)

Other locations where the research will be conducted:
Public elementary schools in Greater Vancouver. Children's homes. During the school year and contingent on the availability of adequate space, testing will be conducted in the child's school. If this is not possible, testing will take place either at the UBC School of Audiology and Speech Sciences, or in the child's home. We have provided the option of testing at the child's home for several reasons. First, this option is provided for parents who would rather their children participate outside of class time. Second, it is possible that the school year will come to an end before testing is completed and we do not want to prevent children from participating. Additionally, for children recruited through our participant database (BREB #B06-0284), some of these individuals do not live within close proximity to UBC, and testing at their home may be more convenient for these individuals.

CO-INVESTIGATOR(S):
Paola Colozzo
Cristy McNiven

SPONSORING AGENCIES:
N/A

PROJECT TITLE:
Who's who? Keeping track of characters during storytelling

REB MEETING DATE:	CERTIFICATE EXPIRY DATE:
February 26, 2009	**February 26, 2010**

DOCUMENTS INCLUDED IN THIS APPROVAL:

DATE APPROVED:
April 22, 2009

Document Name	Version	Date
Consent Forms:		
Consent form recruitment via other third parties	N/A	April 10, 2009
Consent form recruitment via schools	N/A	April 10, 2009
Letter of Initial Contact:		
Information for principals and teachers	N/A	April 10, 2009

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