Referential choice across the lifespan:

Why children and elderly adults produce ambiguous pronouns

Petra Hendriks, Charlotte Koster & John Hoeks

University of Groningen

ABSTRACT

In this study, children, young adults and elderly adults were tested in production and comprehension tasks assessing referential choice. Our aims were 1) to determine whether speakers egocentrically base their referential choice on the preceding linguistic discourse or also take into account the perspective of a hypothetical listener; and 2) whether the possible impact of perspective taking on referential choice changes with increasing age, with its associated changes in cognitive capacity. In the production task, participants described picture-based stories featuring two characters of the same gender, making it necessary to use unambiguous forms; in the comprehension task, participants interpreted potentially ambiguous pronouns at the end of similar orally presented stories. Young adults (age 18-35) were highly sensitive to the informational needs of hypothetical conversational partners in their production and comprehension of referring expressions. In contrast, children (age 4-7) did not take into account possible conversational partners and tended to use pronouns for all given referents, leading to the production of ambiguous pronouns that are unrecoverable for a listener. This was mirrored in the outcome of the comprehension task, where children were insensitive to the shift of discourse topic marked by the speaker. The elderly adults (age 69-87) behaved differently from both young adults and children. They showed a clear sensitivity to the other person's perspective in both production and comprehension, but appeared to lack the necessary cognitive capacities to keep track of the prominence of discourse referents, producing more potentially ambiguous pronouns than young adults, though fewer than children. In conclusion then, referential choice seems to depend on perspective taking in

language, which develops with increasing linguistic experience and cognitive capacity, but also on the ability to keep track of the prominence of discourse referents, which is gradually lost with older age.

ACKNOWLEDGMENTS

INTRODUCTION

A fundamental function of language is reference. This function allows us to talk about the world surrounding us. In particular, it allows us to indicate to others what things in the world we are talking about. What makes reference in language far from trivial is the fact that the correspondence between a linguistic form and its referent in the world is a many-to-many correspondence. Speakers can choose between various forms when referring to a particular referent. Likewise, listeners must often select the intended referent from a range of potential referents. For example, when referring to the queen of the Netherlands, speakers may use the indefinite noun phrase (NP) *a queen*, the definite description *the queen* or the pronoun *she*. Conversely, from the perspective of the listener, referring expressions are highly ambiguous and can refer to various referents, depending on their context of use. A listener may interpret the definite description *the queen* differently when the conversation is about Great Britain than when the conversation is about the Netherlands, and pronouns such as *she* allow for an even wider range of interpretations. So how do speakers decide which form to use and listeners which referent to select?

Apparently, the lexical-semantic content of a referring expression does not fully determine its interpretation and use. Rather, additional factors must come into play that further narrow down the speaker's and listener's choices. This paper is concerned with these factors. Is referential choice mainly restricted by the properties of the linguistic discourse, or must language users also consider the perspective of their conversational partner? And in what way do these linguistic processes depend on general cognitive processes such as those involved in memory?

The goal of this study is to elucidate the processes involved in linguistic reference by investigating the choice of referring expression in discourse by children, young adults and elderly adults. This study stands out from earlier research on referential choice in two aspects. First, because of the design of the linguistic tasks, this study is able to tease apart the effects of the linguistic discourse from the effects of perspective taking in language. Second, this study takes a lifespan approach and compares the performance on the exact same tasks of three age groups differing in linguistic experience and cognitive capacities. As a result, this study allows for a detailed investigation of the interaction between the linguistic and cognitive factors involved in referential choice in narrative discourse.

Factors influencing referential choice

The speaker's choice of referring expression is dependent on the properties of the linguistic and extra-linguistic context. In the linguistic context, as in the visual context, some information is more prominent than other information and hence receives more attention. The prominence of information in the linguistic discourse is referred to by terms such as accessibility, topicality, focus and salience, and is argued to influence how explicit reference should be. Referents that are highly prominent in the discourse can be referred to by short, reduced forms such as unstressed pronouns, whereas referents that are less prominent or new require more explicit forms such as definite or indefinite descriptions (e.g., Givón, 1983; Gundel, Hedberg, & Zacharski, 1993; Grosz, Joshi, & Weinstein, 1995; Arnold, 1998). Discourse properties increasing the referent's prominence include previous mention (Chafe, 1976; Prince, 1992), first mention

(Gernsbacher & Hargreaves, 1988), recency of mention (Givón, 1983), and syntactic

prominence, in particular grammatical subjecthood (Arnold, Eisenband, Brown-Schmidt,

& Trueswell, 2000; Kaiser & Trueswell, 2011).

According to the *discourse-oriented view* on referential choice (e.g. Ariel, 1990;

Givón, 1983), speakers and listeners independently base their referential choices on the

discourse. As the speaker signals the discourse prominence of the referent through the

choice of referring expression, the listener should be able to infer the identity of this

referent from the speaker's choice of referring expression. Others, however, adopting a

listener-oriented view, argue that speakers take into account their listeners in their choice

of referring expression (e.g., Gundel et al., 1993). According to this view, speakers

adhere to Grice's maxim of quantity, which says that speakers should make their

contribution as informative as required, but not more informative than that (Grice, 1975).

As less informative forms tend to be shorter and thus require less effort by the speaker,

speakers aim for such forms and therefore prefer pronouns to full (i.e., non-pronominal)

NPs. However, they will only use a pronoun if the pronoun allows the listener to identify

the intended referent. Thus, speakers do not solely base their choice of form on the

discourse prominence of the referent, but also consider the listener's interpretation of the

form to be used. Speakers must calculate whether this form is informative enough for the

listener in the context of use to be able to identify the intended referent.

Previous studies on referential choice have yielded mixed results regarding the

role of the listener. On the one hand, speakers tend to provide more information if the

intended referent may be difficult to identify for the listener (Arnold & Griffin, 2007;

Francik, 1985). On the other hand, speakers do not circumvent temporary ambiguities in

7

their syntactic choices and thus do not appear to be so concerned with their listeners (Arnold, Wasow, Asudeh & Alrenga, 2004; Ferreira & Dell, 2000). A complicating factor is that it is not always easy to distinguish listener-oriented choices from discourse-oriented choices. Speakers and listeners often have access to the same discourse context. Therefore, what seems to be a listener-oriented choice by the speaker may in fact be a choice based on features of the discourse (Arnold, 2008; Shintel & Keysar, 2009). Distinguishing between discourse-oriented choices and listener-oriented choices requires a detailed consideration of the discourse and the needs of the listener at each point in the discourse. In this study, we manipulate the discourse in such a way that we can tease apart the effects of the discourse context from the effects of perspective taking.

Reference in children and elderly adults

Discourse-oriented processes and listener-oriented processes may tax the language user's cognitive resources differently. A possible way to investigate their effects is by studying referential choice in children and elderly adults. Children have less linguistic experience than adults and their cognitive capacities are still developing. While elderly adults have ample linguistic experience, adult ageing is often characterized by cognitive decline such as limitation of working memory capacity (Carpenter, Miyake, & Just, 1994; Tun, Wingfield, & Stine, 1991) and attention span (Hartley, 1992).

Studies of language production and comprehension in young children have shown that children from the age of 2 or 3 on are already sensitive to the structure of the prior discourse (Campbell, Brooks, & Tomasello, 2000; De Cat, 2011; Hickman & Hendriks,

1999; Song & Fisher, 2005, 2007; Wittek & Tomasello, 2005). Therefore, it is surprising that children acquiring English and Dutch have been found make comprehension errors with object pronouns until as late as age 6 (e.g., Chien & Wexler, 1990; Grimshaw & Rosen, 1990; Koster, 1993; Philip & Coopmans, 1996). Unlike adults, children allow *him* in the sentence *Bert is washing him* to refer to the subject *Bert*. This stands in stark contrast to these same children's production of object pronouns in the same syntactic binding environments, which is adult-like (De Villiers, Cahillane, & Altreuter, 2006; Matthews, Lieven, Theakston, & Tomasello, 2009; Spenader, Smits, & Hendriks, 2009). Thus, these children show a surprising asymmetry in their acquisition of object pronouns, with production preceding comprehension (see Hendriks & Koster, 2010, for discussion of this asymmetry and other asymmetries in child language).

Another aspect of reference that appears to be difficult for children until quite a late age is the correct production of pronouns in discourse. In a large-scale study of children's narrative production, Karmiloff-Smith (1985) found that 4- and 5-year old English and French children would use strings of pronouns to refer at times to the main character of the story and at other times to the other character of the same gender, resulting in ambiguity for the listener.

So, on the one hand, children from an early age on are sensitive to what has been mentioned and how this has been mentioned. On the other hand, children until a relatively late age interpret object pronouns overly generally and overuse pronouns in narratives, resulting in interpretations that were not intended by the speaker and forms that are unrecoverable for a listener. This dichotomy suggests that referential choice is not a unitary phenomenon but involves processes of different kinds.

9

Like children, elderly adults also appear to deviate from young adults in their referential skills. An increased use of pronouns has not only been found in individuals with Alzheimer's disease (Almor et al., 1999), but also in healthy elderly adults. Such ambiguity of reference in healthy adults' language production has been attributed to limitations of memory span (Cohen, 1979; Pratt, Boyes, Robins, & Manchester, 1989). It increases sharply from 75 years old on but likely already emerges in middle-aged individuals (Marini, Boewe, Caltagirone, & Carlomagno, 2005; Ulatowska, Hayashi, Cannito, & Fleming, 1986). Although the language comprehension skills of elderly adults seem to remain relatively intact (Burke, Mackay, & James, 2000), elderly adults are more affected than young adults by intervening linguistic material between a subject pronoun and its antecedent (Light & Capps, 1986). This suggests that elderly adults may be less sensitive to discourse prominence than younger adults.

Asymmetric Grammar Hypothesis

The present study aims to determine whether errors in reference, such as for instance the overuse of pronouns, result from insensitivity to the properties of the linguistic discourse or from difficulty in taking into account the perspective of the listener. We hypothesize that the interaction between discourse prominence and perspective taking is deeply entrenched in the grammar. By the term 'grammar', we refer to all linguistic knowledge involved in speaking and understanding, including knowledge about the discourse. This view of grammar contrasts with a strongly modular view as reflected in most work in

generative syntax since Chomsky (1965) but is compatible with the non-modular view that is prevalent in many other linguistic frameworks (Sag & Wasow, 2011).

We base our predictions on the *Asymmetric Grammar Hypothesis*, according to which the grammar defines a different correspondence between forms and meanings for production than for comprehension (Koster, Hoeks & Hendriks, 2011; see also Blutner, de Hoop & Hendriks, 2006; Hendriks, de Hoop, Krämer, de Swart, & Zwarts, 2010). Such an asymmetric grammar cannot be obtained through reversible rules but requires direction-sensitive constraints. Direction-sensitive constraints are sensitive to their direction of use and only have an effect on the output in production but not in comprehension, or vice versa. An example is the constraint that less informative forms such as pronouns are preferred to explicit forms such as full NPs (Hendriks, Englert, Wubs, & Hoeks, 2008; cf. Gundel et al., 1993). This constraint only affects production, as in comprehension the form to be interpreted is already given. Another constraint, one that is relevant in comprehension, holds that pronouns refer to the most prominent referent in the discourse. Through the use of direction-sensitive constraints, a constraint-based grammar is able to account for asymmetries between production and comprehension in child language, such as the one with object pronouns mentioned above (Hendriks et al., 2010; Hendriks & Spenader, 2005/6; Smolensky, 1996).

To achieve communicative success in spite of the asymmetries between production and comprehension generated by an asymmetric grammar, speakers must take into account the listener's perspective when determining which referring expression to use (Hendriks et al., 2008). Likewise, listeners must take into account the speaker's perspective when deciding on the best interpretation for a referring expression (Hendriks

11

& Spenader, 2005/6; de Hoop & Krämer, 2005/6; van Hout, Harrigan & de Villiers, 2010). In a constraint-based grammar, such perspective taking can be formalized as bidirectional optimization (Blutner, 2000): Language users not only determine the optimal output from their own perspective, but also optimize from the perspective of their conversational partner. Crucially, in this type of grammar-driven perspective taking, speakers and listeners are not concerned with their actual conversational partner but rather with a *hypothetical* conversational partner.

In production, the process of choosing a referring expression can thus be seen as consisting of two steps: First, the speaker chooses the form preferred by the constraints of the grammar in the current discourse. For example, due to the constraint that pronouns are preferred to full NPs, the grammar will select a pronoun for reference to a given referent. In the next step, the speaker calculates whether this form would be interpreted by a hypothetical listener as intended, assuming that this listener uses the same grammar but in the opposite direction (to select an interpretation for the heard form, rather than to select a form for the intended meaning) and has access to the same discourse prominence information as the speaker. If the intended meaning is not recoverable for a listener, the speaker must discard this form and select another form instead. Suppose, for example, that the speaker wishes to refer to a non-prominent referent. If the speaker uses a pronoun, this pronoun will be interpreted by a hypothetical listener as referring to the most prominent referent in the discourse, in accordance with the constraints of the grammar. As this is not the referent intended by the speaker, the speaker must discard the pronoun and use a more explicit full NP.

From this grammar-driven approach to perspective taking, it follows that completing this second step of mentalizing requires additional cognitive resources, as performing two steps is more complex and takes more time than performing only the first of these two steps. In particular, it has been argued that this second step is dependent on sufficient processing speed (van Rij, van Rijn & Hendriks, 2010). If a speaker does not succeed in completing this second step, the selected form will be the form that was chosen on the basis of the grammar in the first step.

We thus predict that less informative and short forms for *highly prominent referents* are easy to produce, as they will already be preferred by the speaker in the first step. Hence, it does not matter whether the speaker succeeds in mentalizing about the listener: The produced form will be the same. On the other hand, more explicit and longer forms for *less prominent referents* are much more difficult to produce, as they require that the less informative form that is preferred in the first step is discarded in the second step, as a result of the speaker's calculation of the listener's interpretation. The grammar prefers pronouns over full NPs even for less prominent referents, so the speaker must block the use of a pronoun for a non-prominent referent through mentalizing.

As this kind of mentalizing requires sufficient cognitive resources, we expect speakers with limited cognitive resources to use more ambiguous pronouns than speakers with greater cognitive resources. Our predictions based on the Asymmetric Grammar Hypothesis are orthogonal to the predictions of discourse-oriented accounts such as put forward by Arnold, Bennetto and Diehl (2009). In their study, it is assumed that pronouns are more difficult to produce than full NPs because pronouns must be properly licensed by the context, whereas full NPs can be used in all contexts. According to this discourse-

oriented account, speakers are expected to avoid using pronouns in cognitively

demanding situations or with limited cognitive resources. In our grammar-driven

perspective taking approach, on the other hand, speakers are predicted to use more

pronouns in such situations, even if this results in ambiguity.

For listeners, as for speakers, the process of interpreting a heard form consists of

two steps. First, the listener determines the best interpretation on the basis of the

constraints of the grammar. Next, the listener considers the perspective of the speaker.

Although explicit forms such as full NPs already provide the necessary information for

identifying the referent in the first step, the second step is not superfluous. Rather,

because pronouns are preferred for reference to the most prominent discourse referent, a

listener will be able to infer from hearing a full NP that its referent must be less

prominent (see also Hendriks et al., 2008). Listeners who do not have sufficient cognitive

resources to complete this second step of mentalizing about the speaker's choices will fail

to notice the decrease in discourse prominence signaled by the speaker by using a full

NP rather than a pronoun. This may affect the listener's decisions later on in the

discourse, for example when interpreting a pronoun.

Although perspective taking has been argued to be cognitively demanding and

therefore not plausible as an online process involved in communication, the proposed

grammar-driven approach to perspective taking assumes that this process can become

less effortful and largely automatic with linguistic experience if the language user has

sufficient memory resources. Computational simulations within the cognitive architecture

ACT-R (Hendriks, van Rijn & Valkenier, 2007; van Rij et al., 2010, 2011) suggest that

this may be a cognitively plausible assumption. If referential choice is dependent on

linguistic experience and cognitive capacity, it becomes relevant to see how the skill of linguistic reference changes across the lifespan.

Methodological approach and predictions

In this study, we investigate in what way speakers' referential choices change within the linguistic discourse and across the lifespan. Children, young adults and elderly adults were presented with four storybooks with pictures that were shown one at a time, and were asked to tell the story to someone else who could not see the pictures. Each storybook featured two characters of the same gender. We analyzed references to the two characters at several positions in the story.

Because the stories are told on the basis of pictures, we can in most cases determine who the intended referent is. Moreover, as the pictures are presented and described one at a time as the story unfolds, rather than as a complete story that has to be retold later, the participants do not have to rely on short-term memory to tell the story. Thus we can eliminate imperfect story recall as a confounding factor. A further advantage of this approach is that we can elicit a particular narrative structure by changing the relative visual prominence of the two characters at particular points in the story. The storybooks are construed in such a way as to elicit a shift of the discourse topic in two positions in the story. Halfway through the story, a first shift is elicited from the character that is introduced first to a second character that is introduced later. At the end of the story, a second shift is elicited from this second character back to the first character.

In this study, we use the term discourse topic to refer to the most prominent referent at a particular point in the discourse. The most prominent discourse referent is defined as the syntactically most prominent referent of the previous sentence that is also referred to in the current sentence. This definition of discourse topic is identical to the definition of backward looking center in Centering Theory (Brennan, Friedman & Pollard, 1987; Grosz et al., 1995; Walker, Joshi & Prince, 1998) and treats the discourse topic as a local notion. That is, the discourse topic is not fixed per story but can shift from one utterance to the other, depending on the linguistic structure of the utterances and the referents mentioned. A topic shift occurs if the discourse topic at the current utterance is different from the discourse topic at the previous utterance.

The design of the stories allows for a detailed analysis of various aspects of reference in discourse. On the basis of the story design, we can formulate predictions about the participants' choices at different moments in the produced narratives.

New referents in a linguistic discourse are generally introduced by an indefinite NP. However, if the referent is sufficiently familiar, it may also be introduced by a definite description. If participants are sensitive to the discourse and know that new referents must be referred to by explicit forms, they will therefore use a non-pronominal (i.e., full) NP to introduce a new character.

Maintained reference to a character is expected to occur by using a definite description, or, if the character is highly prominent at that position in the discourse, a pronoun. To determine its discourse prominence, it is sufficient to know whether the referent has been mentioned before, and if so, how recently it was mentioned and how syntactically prominent it was. Thus the participant may only need to consider the

16

previous discourse when determining the choice of referring expression, without any need for mentalizing about the listener's knowledge.

When the speaker *re-introduces* an earlier-mentioned character who no longer is the topic of the discourse, the situation is different. At this moment, two referents have been introduced into the linguistic discourse that have both been referred to in subsequent utterances. Moreover, another referent than the one to be referred to is the most prominent referent and hence the current discourse topic. If the speaker uses the more economical pronoun to re-introduce an earlier character, this pronoun will be interpreted by the listener as referring to the current discourse topic, and not the re-introduced character. The speaker's use of a pronoun will thus lead to unrecoverability of the intended meaning. Therefore, a speaker who takes into account the listener's perspective will use a full NP at this point in the discourse. So, whereas choosing the appropriate form for introducing and maintaining reference can be based on the properties of the discourse and does not require the speaker to mentalize about the listener's knowledge, re-introducing a referent after a topic shift crucially requires speakers to take into account the listener's perspective.

As children's performance in comprehension may differ from their performance in production, we also administer a similar task of pronoun comprehension in discourse to all three groups. Listeners who are unable to take into account the speaker's referential choices may fail to correctly understand the use of a more explicit expression as signaling a shift in topic. Furthermore, as the three age groups investigated in this study – children, young adults and elderly adults – are often presumed to have different levels of memory

capacity, and memory could be a contributing factor to variance in referential choice, the three groups are also tested on the same memory task.

METHOD

Participants

Participants were children, young adults and elderly adults, all native speakers of Dutch. The youngest group included 31 typically developing children who attended kindergarten classes at a public school (15 girls and 16 boys, mean age 5;6 years, range 4;3-6;5). Two additional children who were tested hardly produced any complete utterances. Therefore, they were left out of the analysis. The young adult group included 20 participants (11 women and 9 men, mean age 26.2 years, range 18-35), students as well as non-students. None of the students were language majors; the non-students had various backgrounds. In the elderly group, there were also 20 participants (10 women and 10 men, mean age 78.7, range 69-87). All the elderly adults were in good health, lived independently and needed a minimum of help in their daily household chores.

Materials production task

The production task was storytelling, based on the pages of a picture storybook. The task began with an introductory page showing all the story characters, followed by a practice storytelling session with two two-page storybooks. The four storybooks used for testing

had six individual picture-pages each, presented to the participants one by one. The storybooks all had the same internal structure designed to elicit topic shifts. An example of a storybook is given in Figure 1.

(FIGURE 1 ABOUT HERE)

Each story featured two characters of the same gender. The first and second pictures showed the first character only. The speaker had to decide how to introduce this character and how to continue referring to it. Because this character was the only possible referent at this point, it was most likely to be the initial topic of the discourse. In the third picture, a second character entered the story. In the next two pictures, this second character was shown performing an action while the first character was passively watching. Again, the speaker had to decide how to introduce this second character and how to continue referring to it. Because the second character had become highly prominent, the speaker was likely to initiate a topic shift to establish the second character as the new topic of the discourse. The final picture once again showed only the first character and the speaker was expected to initiate a topic shift again, switching back to the initial topic, the first character.

Materials comprehension task

The comprehension task included pre-recorded stories about two characters of the same gender. In total, there were eight stories composed of six sentences each. The concluding

sentence of each story contained a potentially ambiguous pronoun, as the gender of the pronoun matched both characters. Four of the stories were designed in such a way that they were parallel in structure to the production stories. These stories included a topic shift in the fourth sentence, marked by changes in form and syntactic position of the two characters. The first character is not referred to anymore by a pronoun but rather by a definite full NP, which appears in object position rather than in subject position. The subject position is taken over by a definite full NP referring to the second character. The other four stories did not include a topic shift half way through the story. Instead, the first character remained the subject throughout the story. See the Appendix for an example of each of the story types.

Following each story, there was a question assessing the reference of the potentially ambiguous pronoun in the concluding sentence. The participant's answer to this question indicated whether the first or the second character was presumed to be the topic at that moment. The stories were presented in two blocks (topic shift vs. non-topic shift) in two orders. An earlier sentence completion task, carried out with other adults, demonstrated that both characters were plausible as the subject of the final sentence. Therefore, any differences in participants' answers to the comprehension question in this task should be attributed to differences in the participants' sensitivity to the discourse structures of the two stories.

Materials memory task

In addition, an auditory memory task was administered to all participants. As the memory task had to be suitable for participants as young as four, we took the memory task from the Schlichting test for language production (Schlichting et al., 1995) and adapted it for adults. This task requires participants to repeat sets of word lists of increasing length and semantic complexity. The adaptation for adults involved adding sets of longer word lists, construed according to the same principles as the original lists. The score on the memory task was the number of successfully repeated lists, with a possible range from 0 to 25. The cut-off moment was after a participant incorrectly repeated two lists in a row or refused to continue.

Procedure

The participants in each group were tested individually in a quiet area, starting with the elicited production task, then the memory task, and finally the comprehension task. The children had a slightly different procedure than the two adult groups. The sessions with children took place at their school and lasted about 20 minutes. A test leader and an assistant were both present. The test leader sat opposite the child and turned the storybook pages (production), read aloud the words (memory), and listened along with the child to the story recordings (comprehension). The assistant, who sat behind a computer screen further away, played the pre-recorded comprehension task stories and noted the responses on the various tasks. During the production task, it was made clear to the child several times that the assistant could not see the pictures, and that the child's

task was to make sure that the assistant also understood what was happening in the stories.

The test leader began the production task by showing the child an introductory page with the characters from all six storybooks and asking the child to name them. In this way, we could be sure that all participants would be equally capable of identifying the characters during their storytelling. Then the test leader told a story based on a two-page storybook, producing one sentence per page, and asked the child to tell a story based on another two-page storybook. After the practice session, the child was once again reminded that the assistant could not see the storybook but wanted to know what was happening and was then asked to describe the four six-page storybooks. Following production, the child was administered the auditory memory task. The session ended with the comprehension task, which included the eight pre-recorded stories, played on the computer by the assistant. After each story, the test leader asked a question about who was referred to in the potentially ambiguous concluding sentence. If children could not answer the comprehension question after they had heard the story for the first time, they were allowed to hear the story again. Between each production or comprehension story, the child was rewarded with a sticker.

The procedure for the adults' sessions was equivalent to the children's sessions, with a few differences. The young adults were tested in various quiet locations and their session lasted roughly 10-15 minutes. The elderly participants were tested in their homes and their actual testing time was roughly 15-20 minutes. Like the children, the young adults and elderly adults also saw an introductory page with all characters at the start of the production task. Only one tester was present for the two adult groups. The adults were

told that later on someone else would have to listen to their recordings and would have to be able to understand their stories without seeing the pictures. One additional instruction was given to the adult groups: They were requested to limit their picture descriptions to one or two sentences per storybook page, since previous testing had shown that some adults tend to tell very long and detailed stories.

Transcription and coding

All the tasks in the test sessions were recorded. After the test sessions, the narratives were transcribed and scored by one researcher. The transcriptions and scoring were then checked for accuracy by a second researcher. The same procedure was followed for the comprehension task and the memory task.

For the production task, a protocol for scoring was developed. Instructions for coding the *discourse topic* were based on the rules pertaining to local discourse coherence and choice of referring expression from Centering Theory (Brennan et al., 1987; Grosz et al., 1995; Walker et al., 1998). In particular, we used the following step-wise instructions: 1) list all referring expressions in each utterance, 2) remove all referring expressions that were not referred to in the prior utterance, 3) if the resulting list contains exactly one pronoun, the referent of the pronoun is the topic, 4) if the resulting list contains no pronouns or more than one pronoun, the expression in the current utterance that was the syntactically most prominent in the prior utterance is the topic. The second step corresponds to the locality of discourse topics ('backward looking centers') in Centering Theory, the third step corresponds to Rule 1 (if any element of the prior

23

utterance is realized by a pronoun in the current utterance, then the backward looking center must be realized by a pronoun also, see Grosz et al., 1995: 214), and the fourth step corresponds to a ranking of discourse entities ('forward looking centers') on the basis of grammatical function ('subject > object(s) > other') (Grosz et al., 1995: 214).

A *topic shift* was coded if the topic at a certain point in the discourse was different from the previous topic. The participants' responses were first scored for whether they shifted the topic from the first to the second character halfway through the story. Two coders (different from the authors) independently scored the transcripts on this point. Across all participant groups, the coders' judgments with respect to the presence or absence of a topic shift agreed for 95.8% of the items. In case of a difference between the two coders, two of the authors made a final decision. Children realized a topic shift (+TS) from the first character to the second character when talking about the third (or sometimes fourth) picture of the storybook 84% of the time (104 out of 124 stories), young adult speakers 99% of the time (79 out of 80 stories), and elderly speakers 90% of the time (72 out of 80 stories). Failure to realize a topic shift was caused by a participant either focusing too strongly on only one character, resulting in no shift, or alternating the two characters without establishing a topic. Only in productions that realize a first topic shift from the first to the second character, is it necessary for a speaker to later re-introduce the first character with a full NP and, therefore, further analyses include only these productions.

In the remaining productions, the first reference to the first character of the story was coded as Intro-1 and the first reference to the second character as Intro-2. The next reference to the first and second character after a reference to this character in the

preceding utterance was coded as Maintain-1 and Maintain-2, respectively. The first reference to the first character as the subject of the sentence after the topic shift halfway through the story was coded as Re-Intro-1. These five references in the discourse were scored for their grammatical forms. Again, two coders independently scored the transcripts on these points.

RESULTS

Production task

To investigate whether the three participant groups differed in their use of full NPs at each of the five positions in the story (i.e., (1) introduction of first character, (2) maintained reference to first character, (3) introduction of second character, (4) maintained reference to second character, and (5) re-introduction of first character), we conducted a Repeated-Measures ANOVA on the arcsine-transformed proportions of full NP use, calculated for participants (F1-analysis) and for items (F2-analysis). Position was considered a within-participants and within-items factor, and Group a between-participants and within-items factor. To guard against possible violations of the statistical assumption of sphericity, the Huynh-Feldt correction was used whenever factors with more than two levels were involved in the analysis (Stevens, 1992). We report the actual degrees of freedom (rounded to the nearest integer) that were used in the statistical test. For ease of exposition, we will present percentages (rounded to the nearest integer)

instead of proportions. Percentages of using a full NP at the five different positions are presented graphically in Figure 2 and numerically in Table 1.

(FIGURE 2 ABOUT HERE)

(TABLE 1 ABOUT HERE)

The results from the statistical analyses showed main effects of Age Group ($F1(2,68)=17.1$; $p<.001$; $F2(1,4)=55.4$; $p<.005$) and of Position ($F1(3,237)=164.2$; $p<.001$; $F2(4,4)=57.5$; $p<.001$). These main effects were, however, qualified by a significant interaction between Age Group and Position ($F1(7,237)=8.2$; $p<.001$; $F2(5,14)=5.5$; $p<.01$). Follow-up analyses were conducted per position. For participant data, One-Way ANOVAs were performed with Age Group as between-participants factor, followed by Bonferroni corrected pairwise comparisons; for item data, Repeated Measures ANOVAs with Age Group as within-items factor were performed, also followed by Bonferroni corrected pairwise comparisons. In reporting the results, we will start with the two introduction positions, then proceed to the two maintenance positions, and end with the final re-introduction position.

Introduction positions: We expected speakers who are sensitive to the linguistic discourse to use referring expressions in accordance with their discourse prominence. When a character is introduced in the story for the first time, a linguistically mature speaker is expected to use a full NP. Age Groups differed significantly at the first and second introduction: Intro-1 ($F1(2,68)=6.6$; $p<.005$; $F2(2,6)=14.3$; $p<.01$) and Intro-2

26

(F1(2,68)=3.3; p<.05; F2(2,6)=9.5; p<.05). At Intro-1, children used significantly fewer full NPs than young adults or elderly adults. At Intro-2, elderly participants used significantly fewer full NPs than young adults; children were numerically between young adults and elderly adults and did not differ significantly from either group. In spite of these group differences for the two introduction moments, all three groups showed roughly the same expected pattern by introducing the first and second characters with a full NP most of the time. Thus, all three groups appear to be sensitive to the given/new status of referents.

Maintenance positions: Speakers who are sensitive to the discourse were predicted to use pronouns to refer to given and highly prominent referents. To investigate this, we looked at how speakers continue referring to a character that has already been introduced in the discourse. At Maintain-1, when maintaining reference to the first character, we found no significant differences between Age Groups. The participants were very homogeneous in their choice of a pronoun. Again, this shows their sensitivity to the given/new status of referents.

In contrast, at Maintain-2 the difference between Age groups was significant (F1(2,68)=18.1; p<.001; F2(2,5)=15.3; p<.01). The young adults no longer preferred a pronoun for maintained reference to the second character, but mostly used full NPs. The elderly adults, in contrast, overwhelmingly used pronouns. The children resembled neither the young adults nor the elderly adults when maintaining reference to the second character. Their choice of form is roughly divided between full NPs and pronouns.

Re-introduction position: Re-introducing a referent after a topic shift was predicted to not only require sensitivity to the linguistic discourse but also perspective

27

taking. We expected speakers who take into account the listener's perspective to produce full NPs at this position in the discourse. In contrast, speakers who are unable to do so are expected to produce ambiguous pronouns. The re-introduction of the first character, Re-Intro-1, showed a significant Age group difference (F1(2,68)=15.9; p<.001; F2(2,5)=27.4; p<.005). At this position at the end of the story, both children and elderly adults used significantly fewer full NPs than young adults. The young adults overwhelmingly re-introduced the first character with a full NP. This shows that the young adults recognized the needs of their listeners. The children and elderly adults, on the other hand, produced many ambiguous pronouns. However, using a pronoun to re-introduce the first character does not necessarily imply that the speaker is insensitive to the listener's needs, as a preference for a pronoun could also be caused by an incorrect estimation of the discourse prominence of the referent.

To see whether the speakers' decisions at the fifth moment in the discourse (at Re-Intro-1, where full NPs are obligatory) differ from the speakers' decisions at the previous moment (at Maintain-2, where full NPs are optional), we ran an additional set of Repeated-Measures ANOVAs, with the factor Position having two levels (Maintain-2 vs. Re-Intro-1). The factor Position was significant by participants but not by items (F1(1,68)=15.0; p<.001; F2(1,3)=2.8; p=.191), and the main effect of Group was significant by participants and by items (F1(2,68)=21.5; p<.001; F2(2,6)=104.9; p<.001). These main effects were qualified by the interaction of Position x Group which was significant by participants, but not by items (F1(2,68)=8.5; p<.001; F2(2,6)=2.2; p=.196). Follow-up analyses showed that children did not change their response between the two positions. For young adults, the number of full NPs increased, which was significant by

participants, but not by items ($F1(1,19)=11.1$; $p<.005$; $F2(1,3)=4.3$; $p=.13$). For the elderly adults, the number of full NPs also increased, and this was significant in both analyses ($F1(1,19)=19.8$; $p<.001$; $F2(1,3)=15.7$; $p<.05$).

Comparing the use of full NPs between the re-introduction moment and the maintained reference moment showed differences between children and elderly adults. Whereas the children used pronouns equally often at the re-introduction moment and the maintained reference moment, the young adults and the elderly adults showed an increase in full NP use at the re-introduction moment. That is, young adults and elderly adults make a difference between a situation in which considering the listener is crucial for obtaining successful communication, and a situation in which it is not. They use more explicit forms when the listener must be taken into account. Children, on the other hand, do not distinguish between these two situations. These results suggest that, whereas young adults and elderly adults are sensitive to the listener's needs, children do not consider the listener.

So children seem to base their referential choices on the discourse prominence of referents only. New referents are referred to by full NPs, and given referents tend to be referred to by pronouns, even if using a pronoun would lead to an incorrect interpretation by the listener. The production results of the young adults and elderly adults, in contrast, suggest that they base their referential choices also on the listener's needs: Their use of pronouns is restricted by the listener's ability to identify the intended referent. However, in these two groups we find two deviations from the expected pattern that full NPs are used for introducing and re-introducing referents and pronouns are used for maintaining reference. First, the young adults use full NPs rather than pronouns for maintained

reference to the second character and thus seem to be overly informative. Second, the elderly adults produce some ambiguous pronouns toward the end of their stories, which is unexplained if they consider the listener's needs. We address these deviations from the expected pattern in the discussion section.

Comprehension task

In the comprehension task, the participants listened to stories with and without a topic shift halfway through the story. The three age groups were compared as to how they answered the question about the potentially ambiguous pronoun at the end of the two types of stories. Response percentages for children, young adults and elderly adults (aggregated over participants) for these two story types are presented in Figure 3 (without topic shift) and Figure 4 (with topic shift).

(FIGURE 3 ABOUT HERE)

(FIGURE 4 ABOUT HERE)

Repeated measures ANOVAs were applied to the response proportions per participant (F1-analysis) and per item (F2-analysis). These analyses included three factors: Response Type (Character 1, Character 2, and Other response), Story Type (with vs. without topic shift) and Age Group (children, young adults, and elderly adults). Response Type was considered a within-participants and within-items factor, Story Type was treated as

within-participants and between-items, and Age Group as between-participants but within-items. As in the previous analyses, an arcsine transformation was used for all proportions. To guard against possible violations of the statistical assumption of sphericity, the Huynh-Feldt correction was used whenever appropriate (Stevens, 1992). We report the actual degrees of freedom that were used in the statistical test, rounded to the nearest integer.

In the comprehension task, we expected to see a distinction between listeners who are, and listeners who are not, sensitive to the speaker's marking of topic shift. The discourse topic at the end of the stories *without* a topic shift is the first character in the story; the discourse topic at the end of the stories *with* a topic shift is the second character that appeared later in the story. Listeners who are sensitive to the speaker's referential choices are predicted to interpret the potentially ambiguous pronoun at the end of the structured discourse as referring to the discourse topic. In contrast, listeners who fail to consider the speaker's choices will allow the ambiguous pronoun to also refer to the other referent in the story.

The main effect of Response Type was significant ($F1(2,135)=108.2$; $p<.001$; $F2(2,12)=38.6$; $p<.001$), as were the two-way interactions of Response Type and Age Group ($F1(4,135)=8.3$; $p<.001$; $F2(4,24)=10.3$; $p<.001$) and of Story Type and Response Type ($F1(2,122)=51.8$; $p<.001$; $F2(2,12)=11.5$; $p<.005$). These effects, however, were qualified by a significant three-way interaction of Response Type, Story Type and Age Group ($F1(4,122)=12.1$; $p<.001$; $F2(4,24)=6.6$; $p<.005$).

Follow-up analyses in each of the two conditions were conducted to investigate the nature of this three-way interaction. In the stories *without topic shift*, the interaction

31

Age Group x Response Type was highly significant ($F1(4,136)=8.8$; $p<.001$; $F2(4,11)=15.3$; $p<.001$). Follow-up analysis showed that children gave significantly fewer correct Character 1 responses than either young adults or elderly adults. The pattern for the Other responses was completely reversed: Children gave significantly more Other responses than either the young adults or the elderly adults.

In the stories *with topic shift*, the interaction between Age Group and Response Type was also significant ($F1(4,136)=11.4$; $p<.001$; $F2(2,7)=6.0$; $p<.05$). In this condition, children produced significantly more incorrect Character 1 responses than young adults or elderly adults and were less inclined to choose the correct Character 2 as the referent for the ambiguous pronoun as compared with young adults ($p1<.001$; $p2=.13$) or elderly adults ($p1<.001$; $p2=.25$). Also, children again gave significantly more Other responses than young adults or elderly adults.

The young adults and the elderly adults usually gave the expected response to the comprehension question for both story types and referred to the discourse topic at that point in the story (i.e., Character 1 in stories without topic shift and Character 2 in stories with topic shift). They had some difficulty detecting the topic shift in the stories with topic shift, which is probably caused by the fact that both referents were plausible as antecedents of the pronoun and the stories were produced with neutral intonation. In natural discourses, non-structural factors such as plausibility of interpretation and prosody provide additional clues for detecting a topic shift. In contrast to the young and elderly adults, the children seemed not to distinguish between the stories with and without topic shift. They answered the comprehension question in both stories by

referring to the first character about half the time and referred to the second character or gave other responses the other half of the time.

So in their comprehension of pronouns in discourse, the elderly adults show the same pattern as the young adults and distinguish between stories with and without topic shift. Their pattern of responses is different from that of the children, who do not distinguish between stories with and without topic shift. Light and Capps (1986) found that elderly adults have no problems disambiguating a pronoun on the basis of information presented in an immediately preceding sentence. Our comprehension results confirm this finding and show that elderly adults disambiguated the pronoun on the basis of the speaker's marking of topic as reflected in the immediately preceding sentence. That is, the elderly adults took into account the syntactic prominence of the potential antecedents in the preceding sentence and chose the syntactically most prominent referent (i.e., the subject of the preceding sentence, which is the discourse topic at that point) as the antecedent of the pronoun.

The comprehension results thus mirror the findings in the production task: Both as speakers and as listeners, elderly adults take into account the perspective of their conversational partner. In contrast, children are insensitive to syntactic prominence. Rather, they seem to base their choice for the antecedent of the pronoun on global aspects of the discourse such as first mention or frequency of use: The first character was the referent that was introduced first and also was the referent that was most frequently referred to in both story types. This would explain children's preference for the first character in both story types. Thus, the sensitivity to subject position found in previous studies with much younger children (Song & Fisher, 2005, 2007) may very well be an

33

effect of global discourse factors such as first mention and frequency of use rather than local discourse factors such as discourse topicality.

Auditory memory task

A participant's score on the memory task was the total number of word lists correctly repeated before he or she made mistakes on two lists in a row. Age Groups differed significantly in mean Memory Score (Univariate ANOVA: $F(2,68)=21.7$; $p<.001$): Young adults had a higher score (mean=11.5; SE=.5) than either children (mean=7.7; SE=.4; $p<.001$) or elderly adults (mean=9.1; SE=.5; $p<.005$). In addition, children tended to have lower scores than the elderly adults ($p=.06$). All pairwise comparisons were Bonferroni corrected.

Age and memory score correlations with language measures

Age and memory score were predicted to be correlated with the outcome measures of the production and comprehension tasks. In particular, we expected higher age and memory score to be positively correlated with the production of full NPs in the production task and with reference to the discourse topic in the comprehension task.

There was no significant correlation between these two predictors (i.e., age and memory score) for either young adults or elderly adults, but there was a marginally significant correlation ($r=.33$; $p=.07$) for children, indicating a trend for memory scores to increase with increasing age.

In relation to the production task, we found significant correlations in children and elderly adults between Age and Memory Score on the one hand, and the proportions of full NPs on the other hand. At *Intro-1*, when the first character is introduced, both children's memory scores and age showed positive correlations with their use of full NPs (r= +.35; p=.05; r= +.62; p<.001, respectively). At *Intro-2*, the moment that the second character is introduced, there was also a significant positive correlation between age and full NP use in children (r= +.43; p<.05). At *Maintain-1*, when maintaining reference to the first character, no correlations between memory or age and the use of full NPs were found in any group. At *Maintain-2*, when maintaining reference to the second character, a significant negative correlation between age and use of full NPs was found in the elderly group: The older the participants, the fewer full NPs were produced (r= -.59; p<.01). At *Re-Intro-1*, when the speaker had to re-introduce the previously mentioned first character, there was a significant positive correlation between memory score and use of full NPs in children: The higher the memory score, the more often a full NP was used (r= +.47; p<.01).

So children's use of full NPs to introduce new referents in the discourse was positively correlated with both higher scores on the memory task and older age. Children's production of full NPs to re-introduce the first character at the end of the story, on the other hand, was positively correlated only with higher scores on the memory task and not with older age. The correlations of the production of full NPs with memory score confirm that producing full NPs is harder for children than producing pronouns. The correlations between age and using a full NP for introducing new referents in the discourse suggest that linguistic experience is an important factor in acquiring the

discourse-oriented aspects of reference. Because no correlation was found between age and the re-introduction of the first character, linguistic experience may not be sufficient for correctly re-introducing given referents. Rather, the ability to block the use of an ambiguous pronoun by taking into account the listener's perspective appears to require additional cognitive resources.

In the elderly adults, a higher age was negatively correlated with full NP use for maintained reference to the second character. That is, the older the participants were, the fewer full NPs they used to continue to talk about this character. So with age, speakers become more economical and produce more pronouns for given referents in multi-referent situations.

In relation to the comprehension task, significant correlations with Age and Memory Score were again found mainly in children and elderly adults. In stories *with a topic shift*, children with higher memory scores were more likely to give the incorrect Character 1 response (r= +.37; p<.05). For the children, Age and the number of incorrect Character 1 answers also showed a strong positive correlation (r= +.53; p<.005). However, a closer look at the data revealed that these unexpected correlations are due to only two observations. For the children, the correlation between Memory Score and Other responses was negative (r= -.56; p<.005), indicating that the higher the memory score, the less likely an irrelevant response was given. Age and proportion of Other responses also showed a negative correlation (r= -.57; p<.005). For the elderly adults, there were two marginally significant correlations: between Memory Score and giving the correct Character 2 response (r= +.41; p=.07) and, inversely, between Memory Score and giving an Other response (r= -.38; p=.10). In the stories *without topic shift*, we only

found a marginally significant – and probably spurious – correlation between Age and giving an Other response (r= +.42; p=.07) for young adults.

So the memory scores of the elderly adults tended to correlate positively with reference to the correct second character in the stories with topic shift. This is in line with the conclusion of the comprehension task that elderly listeners consider the local structure of the discourse. They use the relative syntactic prominence of the two referents in the previous sentence to determine the interpretation of the ambiguous pronoun. Consideration of the syntactic structure of the previous sentence requires access to this structure in memory. The better listeners are able to recall this local structure, the more correct responses they give. The fact that we find this tendency of a positive correlation with memory scores in the elderly adults but not in the young adults suggests that some elderly adults have problems recalling the syntactic structure of the previous sentence.

The negative correlation between children's age and other responses on the comprehension task, and the tendency of a negative correlation between elderly adults' memory scores and other responses, reflect the difficulties these groups have with staying focused on the stories and the referents in the stories. Children and elderly adults with low memory scores would occasionally fail to give a response, refer to characters from previous stories or mention entirely different characters.

DISCUSSION

Considering hypothetical conversational partners

This study proceeded from the Asymmetric Grammar Hypothesis, according to which speakers model a hypothetical listener and listeners model a hypothetical speaker. We found evidence that young and elderly adults take into account hypothetical listeners and speakers when producing and interpreting referring expressions. This allows them to block ambiguous pronouns in production and prevents them from assigning overly general interpretations to pronouns in comprehension. Consequently, they are more explicit as speakers when using a pronoun will lead to misunderstanding by a hypothetical listener. As listeners, they show sensitivity to the speaker's marking of topic shift. Children, on the other hand, do not yet seem capable of taking into account hypothetical listeners and speakers in their production and interpretation of pronouns in discourse. They produce a substantial amount of unrecoverable pronouns and also interpret ambiguous pronouns in a non-adult-like way.

According to the Asymmetric Grammar Hypothesis, the grammar encodes a preference for pronouns over full NPs. Consequently, speakers use pronouns unless there is good reason to use a full NP. That pronouns are the default form is confirmed by our finding that producing full NPs requires a larger memory capacity than producing pronouns. One reason to use a full NP is if the referent is new. In this case, the referent must be properly introduced in the discourse. In our study, the children, young adults and elderly adults all made their referential choices in accordance with this discourse-oriented reason. Another reason to use a full NP is if the listener will not be able to identify the intended referent for the pronoun. This is true if the referent is new, in which case the discourse-oriented reason for using a full NP coincides with the listener-oriented reason. However, the listener may also be unable to identify the antecedent of a pronoun if the

referent has already been introduced in the discourse, but the discourse contains another referent which is a more likely antecedent because it is more prominent at that point in the discourse. The children in our study had difficulty with this purely listener-oriented reason for using a full NP. Our study thus indicates that speakers may be adult-like with respect to discourse-oriented aspects of referential choice while still having difficulty with the listener-oriented aspects of referential choice. This might explain the mixed pattern of results found in previous studies assessing children's referential skills.

In our study, we found two deviations from the expected pattern that full NPs are used for introducing and re-introducing referents and pronouns are used for maintaining reference. First, young adults seemed overly informative in their productions toward the end of the stories, and used full NPs rather than pronouns for maintained reference to the second character. On the basis of its given/new status, we would expect participants to use mainly pronouns here, as they did for maintained reference to the first character. However, there is an important difference between maintaining reference to the first character and maintaining reference to the second character: When the second character has been introduced, the discourse involves two characters of the same gender. It is known that in such situations speakers tend to be more explicit (Arnold & Griffin, 2007; Francik, 1985). It is conceivable that the young adults were more explicit to make it easier for the listener to identify the intended referent. An alternative explanation, proposed by Arnold et al. (2009) to account for the overproduction of full NPs by 9 to 12-year-old children with autism, is that the young adults in our study used full NPs because they had difficulty maintaining activation on the second referent. This explanation seems unlikely, however, as the children and elderly adults in our study would be expected to

have more difficulty maintaining activation on referents but produced fewer full NPs than the young adults.

Whether the use of a full NP for maintained reference to the second character is indeed overly informative depends on how a listener would interpret a pronoun in this position. As the results of the comprehension task show, there is a distinct possibility that the listener does not perceive a speaker's shift in topic. Failing to notice a topic shift will result in misunderstanding a subsequent pronoun. Furthermore, there is considerable variation in the produced narratives. Therefore, in some narratives the speaker may not yet have clearly established the second character as the new topic yet. This may be the case if the second character was not the most prominent referent of the previous utterance because it was mentioned as an object rather than as the grammatical subject. Thus, a speaker using a full NP rather than a pronoun for maintained reference may not be overly informative at all, but is perhaps highly sensitive to the possibility of the listener misunderstanding the pronoun, or alternatively bases the choice of a full NP on the non-prominence of the referent at that point in the narrative.

A second deviation from the expected pattern is the observation that the elderly speakers in our study produced more pronouns than the young adults when the referents had already been introduced. Although elderly speakers show that they are capable of considering the perspective of the listener and produced more full NPs when re-introducing the first character than when maintaining reference to the second character, they produced many more pronouns in both situations than young adults. As the use of a full NP for maintained reference to the second character correlates negatively with age in the elderly adults, elderly speakers may resort to a more economical pronoun when

failing to determine on the basis of the discourse prominence of the referent whether a full NP is necessary. This suggests that elderly speakers have difficulty keeping track of the structure of the discourse and determining the prominence of the referents in the discourse. If this is true, the overuse of pronouns by elderly adults has a different cause than the overuse of pronouns by children.

Considering actual conversational partners

According to the Asymmetric Grammar Hypothesis, the conversational partners modeled within the grammar are hypothetical rather than actual listeners and speakers. Therefore, speakers and their hypothetical listener – and listeners and their hypothetical speaker – base their referential choices on the same discourse information. Indeed, in face-to-face communication, speakers and listeners usually have access to the same visual information. Whereas in other modes of communication, such as telephone conversations but also in the narratives in our study, speakers and listeners do not have access to the same visual information, they do have access to the same prior linguistic discourse. As a consequence of this shared visual or linguistic context, the referential choices of mature language users based on their hypothetical conversational partners will usually be communicatively adequate.

In some discourse situations, however, such as the situations employed in prototypical referential communication tasks (e.g., Barr, 2008; Horton & Keysar, 1996; Lin, Keysar, & Epley, 2010), the information speakers and listeners have access to is crucially distinct, for example because particular information is occluded from the

speaker's view. In such situations, the speaker must monitor which information is available to the speaker but not to the listener and actively inhibit the information that is not available to the listener. This could be an entirely different process. Nevertheless, the pattern Epley, Morewedge and Keysar (2004) found in a referential communication task with adults and children is quite similar to the pattern we found in our study: Adults inhibit an initial egocentric response, whereas children are less capable of doing so and hence behave more egocentrically.

It is conceivable that reasoning about *actual* speakers and listeners is based on reasoning about *hypothetical* speakers and listeners. Considering the perspective of an actual conversational partner presents additional difficulties. Not only must the language user monitor what is shared information and what is privative information, but such perspective taking is also dependent on idiosyncratic properties of conversational partners such as the feedback they give during conversation (Schober, 1993). Considering a hypothetical conversational partner could provide the initial forms and interpretations that are not yet tailored to the actual speaker or listener. These initial forms or interpretations may be adjusted later on the basis of information from the actual communicative situation.

REFERENCES

Ariel, M. (1990). *Accessing Noun-Phrase Antecedents*. London: Routledge.

Almor, A., Kempler, D., MacDonald, M. C., Andersen, E.S., & Tyler, L.K. (1999). Why do Alzheimer patients have difficulty with pronouns? Working Memory, semantics, and reference in comprehension and production in Alzheimer's Disease. *Brain and Language*, 67, 202-227.

Arnold, J. E. (1998). *Reference Form and Discourse Patterns*. PhD Thesis, Stanford University.

Arnold, J. E. (2008). Reference production: Production-internal and addressee-oriented processes. *Language and Cognitive Processes*, 23, 495-527.

Arnold, J.E., Bennetto, L., & Diehl, J.J. (2009). Reference production in young speakers with and without autism: Effects of discourse status and processing constraints. *Cognition*, 110, 131-146.

Arnold, J. E., Eisenband, J.G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The immediate use of gender information: Eyetracking evidence of the time-course of pronoun resolution. *Cognition*, 76, B13-26.

Arnold, J. E., & Griffin, Z. M. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, 56, 521-536.

Arnold, J. E., Wasow, T., Asudeh, A., & Alrenga, P. (2004). Avoiding attachment ambiguities: The role of constituent ordering. *Journal of Memory and Language*, 51, 55-70.

Koster, C., Hoeks, J., & Hendriks, P. (2011). Comprehension and production of subject pronouns: Evidence for the asymmetry of grammar. In A. Grimm, A. Müller, C. Hamann, & E. Ruigendijk (Eds), *Production-comprehension asymmetries in child language* (pp. 99-122). De Gruyter Mouton, Berlin.

Barr, D. J. (2008). Pragmatic expectations and linguistic evidence: Listeners anticipate but do not integrate common ground. *Cognition*, 109, 18-40.

Blutner, R. (2000). Some aspects of optimality in natural language interpretation. *Journal of Semantics*, 17, 189-216.

Blutner, R., de Hoop, H., & Hendriks, P. (2006). *Optimal Communication*. Stanford, CA: CSLI Publications.

Brennan, S. E., Friedman, M. W., & Pollard, C. J. (1987). A Centering approach to pronouns. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics* (ACL '87). Stroudsburg, PA: Association for Computational Linguistics.

Burke, D. M., MacKay, D. G., & James, L.E. (2000). Theoretical approaches to language and aging. In T. Perfect & E. Maylor (Eds.), *Models of Cognitive Aging* (pp. 204-237). Oxford, U.K.: Oxford University Press.

Campbell, A. L., Brooks, P., & Tomasello, M. (2000). Factors affecting young children's use of pronouns as referring expressions. *Journal of Speech, Language, and Hearing Research*, 43, 1337-1349.

Carpenter, P. A., Miyake, A., & Just, M. A. (1994). Working memory constraints in comprehension: Evidence from individual differences, aphasia, and aging. In M.

Gernsbacher (Ed.), *Handbook of Psycholinguistics*. San Diego, CA: Academic
Press.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point
of view. In C. N. Li (Ed.), *Subject and Topic* (pp. 25-56). New York: Academic
Press.

Chien, Y.-C., & Wexler, K. (1990). Children's knowledge of locality conditions in
binding as evidence for the modularity of syntax and pragmatics. *Language
Acquisition*, 1, 225–295.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

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

De Cat, C. (2011). Information tracking and encoding in early L1: linguistic competence
vs. cognitive limitations. *Journal of Child Language*, 38, 828-860.

De Hoop, H., & Krämer, I. (2005/6). Children's optimal interpretations of indefinite
subjects and objects. *Language Acquisition*, 13, 103-123.

De Villiers, J., Cahillane, J., & Altreuter, E. (2006). What can production reveal about
Principle B? In K. U. Deen, J. Nomura, B. Schulz, & B. D. Schwartz (Eds.), *The
Proceedings of the Inaugural Conference on Generative Approaches to Language
Acquisition–North America*, Vol. 1 (pp. 89-100). Honolulu, HI: University of
Connecticut Occasional Papers in Linguistics 4.

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and
adults: Equivalent egocentrism but differential correction. *Journal of
Experimental Social Psychology*, 40, 760-768.

Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, 40, 296-340.

Francik, E. P. (1985). *Referential Choice and Focus of Attention in Narratives*. PhD Thesis, Stanford University.

Gernsbacher, M.A., & Hargreaves, D.J. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language*, 27, 699–717.

Givón, T. (1983). Topic and continuity in discourse: An introduction. In T. Givón (Ed.), *Topic Continuity in Discourse: A Quantitative Cross-language Study* (pp. 1-42). Amsterdam: John Benjamins Publishing.

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

Grimshaw, J., & Rosen, S. T. (1990). Knowledge and obedience: The developmental status of the Binding Theory. *Linguistic Inquiry*, 21, 187-222.

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A framework for modeling the logical coherence of discourse. *Computational Linguistics*, 21, 203-225.

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

Hartley, A. A. (1992). Attention. In F. I. M. Craik, & T. A. Salthouse (Eds.), *The Handbook of Aging and Cognition* (pp. 3-49). Hillsdale, NJ: Lawrence Erlbaum.

Hendriks, P., de Hoop, H., Krämer, I., de Swart, H., & Zwarts, J. (2010). *Conflicts in Interpretation*. London: Equinox Publishing.

Hendriks, P., Englert, C., Wubs, E., & Hoeks, J. (2008). Age differences in adults' use of referring expressions. *Journal of Logic, Language and Information*, 17, 443-466.

Hendriks, P., & Koster, C. (2010). Production/comprehension asymmetries in language

    acquisition. Introduction to special issue on asymmetries in language acquisition.

    Lingua, 120, 1887-1897.

Hendriks, P., van Rijn, H., & Valkenier, B. (2007). Learning to reason about speakers'

    alternatives in sentence comprehension: A computational account. *Lingua*, 117,

    1879-1896.

Hendriks, P., & Spenader, J. (2005/6). When production precedes comprehension: An

    optimization approach to the acquisition of pronouns. *Language Acquisition*, 13,

    319-348.

Hickmann, M., & Hendriks, H. (1999). Cohesion and anaphora in children's narratives: A

    comparison of English, French, German and Mandarin Chinese. *Journal of Child*

    Language, 26, 291-452.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common

    ground? *Cognition*, 59, 91-117.

Kaiser, E., & Trueswell, J.C. (2011). Investigating the interpretation of pronouns and

    demonstratives in Finnish: Going beyond salience. In E. Gibson, & N.

    Pearlmutter (Eds.), *The Processing and Acquisition of Reference* (pp. 323-353).

    Cambridge, MA: MIT Press.

Karmiloff-Smith, A. (1985). Language and cognitive processes from a developmental

    perspective. *Language and Cognitive Processes*, 1, 61-85.

Koster, C. (1993). *Errors in Anaphora Acquisition*. PhD Thesis, Utrecht University.

Light, L. L., & Capps, J. L. (1986). Comprehension of pronouns in young and older

    adults. *Developmental Psychology*, 22, 580-585.

Lin, S., Keysar, B., & Epley, N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology*, 46, 551-556.

Marini, A., Boewe, A., Caltagirone, C., & Carlomagno, S. (2005). Age-related differences in the production of textual descriptions. *Journal of Psychological Research*, 34, 439-463.

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2009). Pronoun co-referencing errors: Challenges for generativist and usage-based accounts. *Cognitive Linguistics*, 20, 599-626.

Philip, W., & Coopmans, P. (1996). The Double Dutch Delay of Principle B Effect. In A. Stringfellow, D. Cahana-Amitay, E. Hughes, & A. Zukowski (Eds.), *Proceedings of the 20th Boston University Conference on Language Development (BUCLD)* (pp. 576-587). Somerville, MA: Cascadilla Press.

Pratt, M. W., Boyes, C., Robins, S., & Manchester, J. (1989). Telling tales: Aging, working memory, and the narrative cohesion of story retellings. *Developmental Psychology*, 2, 628-635.

Prince, E. (1992). The ZPG letter: Subjects, definiteness, and information-status. In W. Mann, & S. Thompson (Eds.), *Discourse Description: Diverse Linguistic Analyses of a Fund-raising Text* (pp. 295-326). Amsterdam: Benjamins.

Sag, I. A., & Wasow, T. (2011). Performance-Compatible Competence Grammar. In R. Borsley, & K. Borjars (Eds.), *Non-Transformational Syntax: Formal and Explicit Models of Grammar*. Oxford: Wiley-Blackwell.

Schlichting, J.E.P.T., van Eldijk, M.C.M., Lutje Spelberg, H.C., van der Meulen, S., & van der Meulen, B.F. (1995). *Schlichting Test voor Taalproductie*. Nijmegen: Berkhout.

Schober, Michael F. (1993). Spatial perspective-taking in conversation. *Cognition*, 47, 1-24.

Shintel, H., & Keysar, B. (2009). Less is more: A minimalist account of joint action in communication. *Topics in Cognitive Science*, 1, 260-273.

Smolensky, P. (1996). On the comprehension/production dilemma in child language. *Linguistic Inquiry*, 27, 720-731.

Song, H., & Fisher, C. (2005). Who's "she"? Discourse prominence influences preschoolers' comprehension of pronouns. *Journal of Memory and Language*, 52, 29-57.

Song, H., & Fisher, C. (2007). Discourse prominence effects on 2.5-year-old children's interpretation of pronouns. *Lingua*, 117, 1959-1987.

Spenader, J., Smits, E. J., & Hendriks, P. (2009). Coherent discourse solves the Pronoun Interpretation Problem. *Journal of Child Language,* 36, 23-52.

Stevens, J. (1992). *Applied Multivariate Statistics for the Social Sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Tun, P. A., Wingfield, A., & Stine, E. A. (1991). Speech-processing capacity in younger and older adults: A dual-task study. *Psychology and Aging*, 6, 3-9.

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

Van Hout, A., Harrigan, K., & de Villiers, J. (2010). Asymmetries in the acquisition of definite and indefinite NPs. *Lingua*, 120, 1973-1990.

Van Rij, J., van Rijn, H., & Hendriks, P. (2010). Cognitive architectures and language acquisition: A case study in pronoun comprehension. *Journal of Child Language*, 37, 731-766.

Van Rij, J., van Rijn, H., & Hendriks, P. (2011). WM load influences the interpretation of referring expressions. In F. Keller, & D. Reitter (Eds.), *Proceedings of the 2nd workshop on Cognitive Modeling and Computational Linguistics* (pp. 67-75). Portland, OR: Association for Computational Linguistics.

Walker, M. A., Joshi, A. K., & Prince, E. F. (1998). *Centering Theory in Discourse*. Oxford: Clarendon Press.

Wittek, A., & Tomasello, M. (2005). Young children's sensitivity to listener knowledge and perceptual context in choosing referring expressions. *Applied Psycholinguistics*, 26, 541-558.

Example of a comprehension story with topic shift

1 Een schoonmaakster wil de eendjes gaan voeren.

   'a cleaning-lady wants to go feed the ducks'

2 Ze haalt het oude brood uit de broodtrommel.

   'she gets the old bread out if the breadbox'

3 Ze vraagt aan een juf om mee te gaan.

   'she asks the teacher(FEM) to go along'

4 De juf scheurt de broodjes van de schoonmaakster in stukjes.

   'the teacher(FEM) tears the cleaning-lady's bread into pieces'

5 En dan geeft de juf het brood van de schoonmaakster aan de eendjes.

   'and then the teacher(FEM) gives the cleaning-lady's bread to the ducks'

6 Ze vindt eendjes hele lieve diertjes.

   'she thinks ducks are very sweet animals'

Comprehension question:

   Wie vindt eendjes hele lieve diertjes?

   'who thinks ducks are very sweet animals?'

Example of a comprehension story without topic shift

1 Een clown heeft net zijn eigen gezicht geschminkt.

'a clown has just painted his own face'

2 Hij wil wel eens iemand anders schminken.

'he wants to paint someone else'

3 Hij komt in de keuken een kok tegen.

'he comes across the cook(MASC) in the kitchen'

4 De clown besluit de kok te schminken.

'the clown decides to paint the cook(MASC)'

5 En dan schminkt de clown een heel stoer gezicht bij de kok.

'and then the clown paints a real tough face on the cook(MASC)'

6 Hij vindt dat het prachtig is geworden.

'he thinks it turned out great'

Comprehension question:

Wie vindt het prachtig geworden?

'who thinks it turned out great?'

Examples of elicited narratives (available through CHILDES as part of the Asymmetries database)

Child (c06, female, age 6:2, topic shift):

Picture 1:   een Pietpiraat met de voetbal.

'a Pete-pirate with the football'

Picture 2:   dan schopt ie [: hij] (he)m.

'then he kicks it(MASC)'

Picture 3: dan is die in het water.

'then it(DEM) is in the water'

Picture 4: dan gaat de ridder (he)m vangen.

'then the knight goes to catch it(MASC)'

Picture 5: en <hij heeft de v> [//] hij heeft de bal in een net gevangen.

'and he has the f- he has caught the ball in a net'

Picture 6: nu heeft ie [: hij] ze [: zijn] bal weer terug.

'now he has his ball back again'

Child (c08, male, age 5;6, no topic shift):

Picture 1: de piraat gaat met de bal spelen.

'the pirate is going to play with the ball'

Picture 2: en toen gooide die in (he)t water.

'and then threw it(DEM) in the water'

Picture 3: en toen begon ie [: hij] te huilen. en toen kwam de ridder (he)m pakken.

'and then he began to cry. and then the knight came to get it(MASC)'

Picture 4: en <toen was de> [/] toen was de piraat zo blij.

'and then the then the pirate was so happy'

Picture 5: en toen ging ie [: hij] (he)m uit het water halen.

'and then he got it(MASC) out of the water'

Picture 6: en toen was ie [: hij] heel blij.

'and then he was very happy'

Young adult (a16, male, age 27, topic shift):

Picture 1:   de piraat met een houten been heeft een voetbal.

'the pirate with a wooden leg has a football'

Picture 2:   hij schopt met z(ij)n houten been de voetbal in de vijver.

'he kicks the football with his wooden leg into the pond'

Picture 3:   en huilt. want hij kan niet meer bij de bal kome(n). de ridder die ziet dat

allemaal.

'and cries. because he can't reach the ball anymore. the knight he(DEM) sees

all that'

Picture 4:   de ridder die pakt een vangnet.

'the knight he(DEM) gets a net'

Picture 5:   en haalt zo de bal uit het water voor de piraat.

'and gets the ball out of the water for the pirate'

Picture 6:   de piraat die heeft een dikke glimlach. want die is blij dat ie [: hij] de bal weer

heeft.

'the pirate he(DEM) has a big smile. because he(DEM) is happy that he has

the ball back again'

Elderly adult (e38, female, age 70, topic shift):

Picture 1:   daar hebbe(n) we de piraat. en die piraat gaat met de voetbal op stap. dus

heeft vast slechte bedoelingen.

'there we have the pirate. and that pirate is going out with the football. so

probably has bad intentions'

Picture 2:   hij schopt de voetbal in (h)et water. en dan is die dus weg.

'he kicks the football in the water. and then it(DEM) is gone'

Picture 3:   en dan komt dus de ridder dr [: er] aan. en die wil hem waarschijnlijk helpen. hij heeft het harnas al aan. dus hij zou misschien zo in (h)et water kunnen om de bal terug te pakken.

'and then the knight arrives. and he(DEM) probably wants to help him. he already has the armor on. so maybe he could go straight into the water to get the ball back'

Picture 4:   oh hij is nog loze [: slim/weg (dialect)] geweest. heeft een schepnet gehaald (.) om de bal uit (h)et water te halen.

'oh he has been smart/away. has picked up a net (.) to get the ball out of the water'

Picture 5:   en hij heeft (he)m dus gepakt. nu heeft dus de [///] &eh de piraat heeft dus de bal terug tenminste als hij hem terug geeft.

'and he has got it(MASC). now has the eh the pirate has the ball back at least if he gives it(MASC) back'

Picture 6:   ja daar is ie [: hij] blij en gelukkig. hij heeft z(ij)n bal weer terug. en nu oppassen dat ie [: hij] niet weer in (h)et water komt.

'yes there he is happy and glad. he has his ball back again. and now watch out that it(MASC) doesn't go in the water again'