Complex language, complex thought? The relation between children's production of double embeddings and Theory of Mind

Bogaerds-Hazenberg, Suzanne; Hendriks, Petra

Published in:
Linguistics in the Netherlands

Publication date:
2016

Citation for published version (APA):
Bogaerds-Hazenberg, S., & Hendriks, P. (2016). Complex language, complex thought? The relation between children's production of double embeddings and Theory of Mind. *Linguistics in the Netherlands*, *33*, 28-40.

Complex language, complex thought? The relation between children's production of double embeddings and Theory of Mind

Suzanne Bogaerds-Hazenberg & Petra Hendriks

University of Groningen

It has been argued (e.g., by De Villiers and colleagues) that the acquisition of sentence embedding is necessary for the development of first-order Theory of Mind (ToM): the ability to attribute beliefs to others. This raises the question whether the acquisition of double embedded sentences is related to, and perhaps even necessary for, the development of second-order ToM: the ability to attribute beliefs about beliefs to others. This study tested 55 children (aged 7-10) on their ToM understanding in a false-belief task and on their elicited production of sentence embeddings. We found that second-order ToM passers produced mainly double embeddings, whereas first-order ToM passers produced mainly single embeddings. Furthermore, a better performance on second-order ToM predicted a higher rate of double embeddings and a lower rate of single embeddings in the production task. We conclude that children's ability to produce double embeddings is related to their development of second-order ToM.

Key words: Language acquisition, false-belief task, sentence embedding, Theory of Mind

1. *Introduction*

Theory of Mind (ToM), or the ability to attribute mental states such as beliefs, desires and intentions to oneself and others, develops during childhood. Children gradually learn that "people live their lives in a mental world as much as in a world of real situations and occurrences" (Wellman, Cross and Watson 2001:656). This enables them to represent other people's thoughts in their own brain, which is needed to predict other people's behavior.

A frequently used task to test ToM abilities is a *false-belief* (FB) task (Wellman, Cross and Watson 2001; Wimmer and Perner 1983). In an FB task, the true knowledge of a participant is pitted against the false belief of a character in the task, for example a puppet. If participants are able to attribute a false belief, contrasting with their own true belief, to the puppet and predict its behavior accordingly, participants are said to possess a first-order ToM. People possess a second-order ToM if they are able to attribute a belief about a belief (that is, a second-order belief) to someone else.

Many studies found strong correlations between first-order FB reasoning and a wide variety of linguistic tasks, suggesting that ToM development and language acquisition are fundamentally related and interdependent (Astington and Baird 2005). Some researchers claim that language acquisition is dependent on ToM development, others that there is a bidirectional relationship between language acquisition and ToM development (De Mulder 2011), and yet other researchers argue that particular linguistic knowledge is necessary for the development of ToM (De Villiers and Pyers 2002). As many FB tasks are verbal in nature, it is conceivable that general language abilities are useful or even necessary to pass an FB task. However, De Villiers and Pyers (2002) argue that there is a specific linguistic prerequisite for ToM, namely the acquisition of the syntax for embedded complements under mental-state verbs. That is, the ability to understand other minds is dependent on the ability to understand and produce sentences like the following:

(1) Mickey thinks that Donald is ill.

This sentence contains the mental-state verb *thinks*. The sentential complement *Donald is ill* is embedded under this verb. To understand this sentence, a child needs to know the lexical meaning of *thinks* and also needs to understand that the proposition embedded under *thinks* can be true or false, as it reports something in Mickey's mind and not in reality. The embedded structure 'organizes' meaning in the mind, and so the acquisition of the language of these propositional attitudes would open up a "classification into worlds, or different points of view on reality" (Jill de Villiers 2005:188). A child must acquire this complex syntactic structure not only for reporting about false beliefs in his language, but also for representing in his own mind the belief states of others. A longitudinal study by De Villiers and Pyers (2002) found that children's spontaneous production of sentential complements is a precursor of FB understanding. Other studies showed that both oral and signing deaf children experiencing specific problems with these sentential complements are also delayed in passing (non)verbal FB tasks (Jill de Villiers 2005; Peter de Villiers 2005). Therefore, the mastery of this piece of syntax is claimed to be crucial to successful first order FB reasoning: without sentential complementation, a child has no means to represent its own beliefs and compare them to other people's beliefs.

Although this language-first hypothesis has initiated many longitudinal and intervention studies on the link between sentence embedding and FB reasoning in various populations (Peter de Villiers 2005), few studies separate first-order FB, which typically develops around age 3 or 4 (Wimmer and Perner 1983), from second-order FB, which has been argued to develop between the ages of 7-9 (Hollebrandse and Van Hout 2015). If it is true that the syntax of single embedding is related to, and perhaps even necessary for, successful first-order FB reasoning, then it can be hypothesized by the same reasoning that the syntax of double embedding (as in (2)) is related to second-order FB reasoning (cf. De Villiers, Hobbs and Hollebrandse 2014).

(2) Daisy thinks that Mickey thinks that Donald is ill.

Only recently, some studies have started to examine whether the language-first hypothesis extends to more complex forms of FB reasoning. Their rationale is similar, namely that "the recursive linguistic representations involved in syntactic embedding may provide the scaffolding to perform the recursive step of a second-order false-belief reasoning" (Hollebrandse and Van Hout 2015:104). Although non-ceiling performance was found for children between the ages of 6 and 10 on tasks testing their comprehension and production of double embeddings as well as on tasks assessing second-order FB understanding (Hollebrandse, Hobbs, De Villiers and Roeper 2008; Hollebrandse and Van Hout 2015), these studies did not look into the relation between children's performance on sentence embedding and the same children's performance on FB reasoning, but only analyzed children's performance on the tasks separately .

Most studies so far used comprehension tasks in which children had to provide an answer to a double-embedded question (e.g., *What did the policeman say the woman said was in her backpack?*). A weakness of such a design is that the target response is not only dependent on the child's mastery of double embedding, but also on its mastery of long-distance dependencies and its ability to deal with both complexities at the same time. Some studies used a production task, in which however it sufficed to repeat a previously mentioned double embedded sentence or only part of it (e.g., Hollebrandse and Van Hout 2015). None of these tasks required the child to construct a complete double embedded sentence. Therefore, it is still unclear to what extent children are actually able to represent this recursive sentence structure and how this ability relates to their ToM development.

The current study aims to fill this gap by investigating the relation between the construction of double embeddings and second-order FB reasoning. It reports on the results of an experiment with Dutch children who were presented with a second-order FB task and a novel production task eliciting double embeddings on the basis of visual stimuli.

2. *Methodology*

2.1 *Participants*

55 typically developing Dutch children (27 boys, 28 girls) from two primary schools participated in the experiment, with a mean age of 8.8 years old (age range 7;6-10;4). The schools provided standardized scores for reading fluency (DMT; a three-minutes word reading test) and reading level (AVI; a measure reflecting the text complexity a child can handle). The scores on these two reading tasks may provide an (albeit very indirect) measure of general language abilities.

2.2 *Materials and design*

All children received a verbal FB task (the Bake Sale task adapted from Hollebrandse et al. 2014) and a novel sentence elicitation task eliciting double embeddings.

Bake Sale task – This verbal FB task consists of eight short stories with corresponding questions measuring first-order (FB1) and second-order (FB2) false-belief. In all stories, two protagonists initially have the same belief, but during the story independently change their beliefs, ending up with different beliefs. When the child is asked a question about the belief of one protagonist about the belief of the other protagonist, this requires second-order FB reasoning. Each story is illustrated by four pictures that focus the child's attention. None of the stories contain syntactic embeddings. The questions only contain single embeddings.

Sentence Embedding task – This production task elicits double embeddings on the basis of pictures. Thus, the task contrasts with the highly verbal task of Hollebrandse et al. (2008) (see also Hollebrandse and Van Hout 2015), which required participants to first interpret a double embedding in a story and then to reproduce the embedded proposition in response to a question. In our task, each picture depicts a character producing a talking cloud (or thought bubble) in which another character is embedded producing another talking cloud (or thought bubble). The embedded talking cloud is filled with two short phrases, on the basis of which a sentence can be produced. Only phrases with an AVI level (a readability level based on word frequency and length, SLO 2008) appropriate for 7-year-olds were used. In a pretest with 8 adults and 2 children, it was checked that formulating simple sentences on the basis of the two phrases did not result in any difficulty. The task consisted of 36 target items (pictures targeting double embedding such as *x says/thinks that y says/thinks that p*), preceded by 12 practice items (pictures targeting single embeddings such as *x says/thinks that p*), as illustrated in Figure 1. The practice items were included for instruction and to ensure that the children understood the task. The content of the practice item on the left can be verbalized as in (3a), whereas the content of the target item on the right can be phrased with a double embedding (3b).

(3) a. Katrien zegt dat Donald naar de slager is.

‘Daisy says that Donald went to the butcher’s.’

b. Mickey denkt dat Katrien zegt dat Donald naar de slager is.

‘Mickey thinks that Daisy says that Donald went to the butcher’s’.

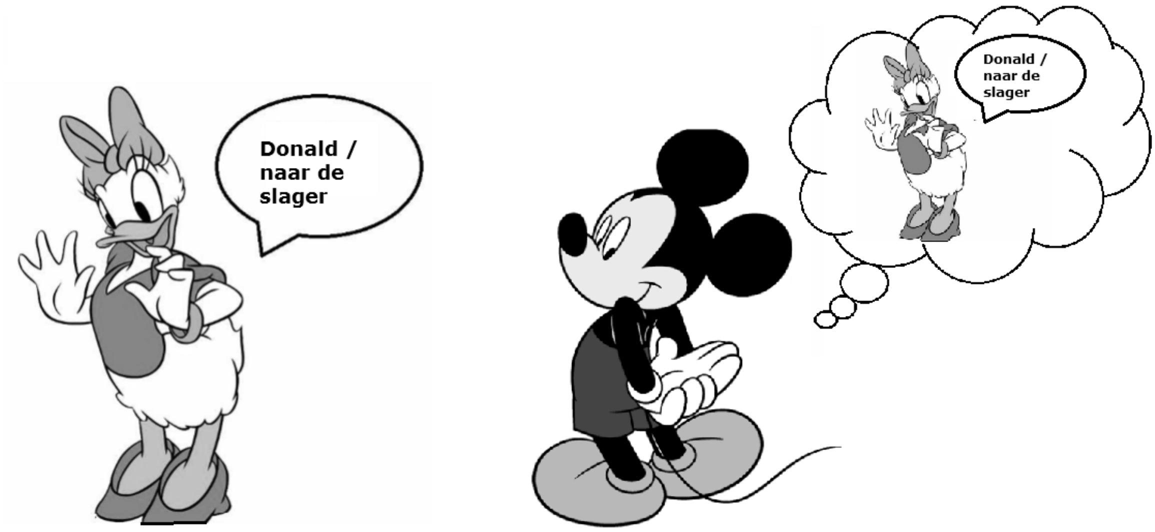

Figure 1. Practice item (left) and target item (right) in the Sentence Embedding task

The distribution of characters, talking clouds and thought bubbles was counterbalanced across items.

2.3 *Procedure*

All children were tested individually by the same experimenter in a quiet room at their school. Each test session took approximately 35 minutes. The order in which the two tasks were conducted was balanced across participants in order to avoid priming effects. Children’s responses were audio-taped for later transcription. The children were told that they could stop the experiment at any moment. At the end of the test session, the child was given a small reward.

The stories of the Bake Sale task were read aloud by the experimenter. At specific points during the story, the FB questions were asked. The children were encouraged to formulate a response.

The Sentence Embedding task started with an introductory story, explaining that an expensive painting had been stolen and that the task of the child was to help the detective find the thief by providing information about what the characters on the pictures said and thought about each other. They were instructed to do so in only one sentence. We did not explicitly mention that the characters could be lying , but the narrative context suggests this possibility. As a result, the embedded complement could be false. De Villiers et al. (2014) argue that contrasts in truth value between the matrix clause and the embedded clause may help the child in recognizing recursive structures. This could also be true for producing recursive structures.

At the beginning of the task, the experimenter checked whether the names of the three characters (Donald, Mickey or Daisy) were familiar to the children and whether the children knew that thought bubbles and talking clouds could be verbalized as *to think* and *to say,* respectively. Then the experiment continued with 12 practice items, which contained only single thought bubbles and talking clouds. The experimenter encouraged the child to produce syntactic embeddings to verbalize the content of the picture, by occasionally pointing out features of the pictures (e.g., “Don’t forget to mention the character” or “Look again, is Mickey thinking or saying something?”) or by asking the child to finish the sentence. Children who spontaneously used syntactic embeddings were praised. Children who omitted information or did not produce syntactic embeddings were told how to phrase the information correctly. Next, the task continued with the 36 target items. From this moment on, the experimenter no longer corrected the child. The experimenter told the child that the pictures would become slightly more difficult, but that the child had done very well and could continue to tell what was on the picture, again in one sentence. After each set of twelve target items, there was a small break so that the child could relax for a minute and stay concentrated.

2.4 *Data analysis of the Sentence Embedding task*

As shown in Table 1, the responses on the Sentence Embedding task were transcribed and categorized into four categories: (1) Double embedding, (2) Mixed embedding, (3) Single embedding and (4) No embedding. Double embedding is the target category and involves a proposition embedded under two verbs. Mixed embeddings involve a combination of indirect speech and direct speech. Indirect speech requires syntactic embedding of the reported proposition, which is marked in Dutch by the complementizer *dat* ('that') and the verb-final word order indicative of subordinate clauses. Direct speech, in contrast, does not involve syntactic embedding and hence is not introduced by a complementizer and has the verb-second word order indicative of Dutch main clauses. In single embedded sentences, typically one of the two characters on the picture is omitted or the two characters are referred to in a DP conjunction. In the case of no embedding, only direct speech is used or sentence embedding is avoided in some other way.

Table 1. Categorization of production data in the Sentence Embedding task

Category	**Example**

Double embedding	*Mickey denkt dat Katrien zegt dat Donald naar de slager is.* Mickey thinks that Daisy says that Donald went to the butcher's.
Mixed embedding	*Mickey denkt dat Katrien zegt: Donald is naar de slager.* Mickey thinks that Daisy says: Donald went to the butcher's. *Mickey denkt: Katrien zegt dat Donald naar de slager is.* Mickey thinks: Daisy says that Donald went to the butcher's.
Single embedding	*Mickey denkt dat Katrien Donald meeneemt naar de slager.* Mickey thinks that Daisy takes Donald to the butcher's. *Mickey denkt dat Donald/Katrien naar de slager is.* Mickey thinks that Donald/Daisy went to the butcher's. *Mickey denkt dat Katrien en Donald naar de slager zijn.* Mickey thinks that Donald and Daisy went to the butcher's. *Mickey en Katrien zeggen/denken dat Donald naar de slager is.* Mickey and Daisy say/think that Donald went to the butcher's.
No embedding	*Mickey zegt: "Katrien, ga jij naar de slager?" "Nee, maar Donald wel".* Mickey says: "Daisy, are you going to the butcher's?" "No, but Donald is". *Mickey (en Katrien/Donald) gaan naar de slager.* Mickey (and Daisy/Donald) go to the butcher's.

3. *Results*

3.1 *False-belief reasoning*

In total, 92% of the FB1 questions were answered correctly by the 7 to 10-year-old children, and 47% of the FB2 questions. We grouped the children according to their performance on the Bake Sale task. FB2-passers answered the FB2 and FB1 questions correctly, FB1-passers answered the FB1 questions correctly but made too many errors on the FB2 questions, and FB-failers made too many errors on both the FB2 and FB1 questions. The threshold for passing FB1 was set at answering at least seven out of the eight FB1 questions correctly. Likewise, the threshold for passing FB2 was set at answering at least seven out of the eight FB2 questions correctly. 12 children (22%) fell into the group of FB2-passers, 37 children (67%) were FB1-passers and 6 children (11%) were FB-failers. Table 2 lists these three FB groups. The mean age of the participants did not differ significantly across groups.

Table 2. Distinction of three FB groups on the basis of the children's performance on the Bake Sale task. Per FB group, mean percentage correct on FB1 and FB2 and mean chronological age are listed (standard error).

	FB2-passers	**FB1-passers**	**FB-failers**	**Total**
FB1	95.8 (2.5)	95.3 (1.0)	62.6 (6.5)	91.8 (1.8)
FB2	91.7 (1.8)	35.8 (4.8)	27.1 (12.3)	47.1 (4.8)
Age	9.2 (0.1)	8.7 (0.1)	8.8 (0.5)	8.8 (0.1)

3.2 *Production of double embeddings*

Based on performance on the practice items, one child (age 7;6) seemed not to understand the task, as this child did not produce any single embeddings on the practice items, in contrast to the other children.[i] The responses of the 55 children on the target items of the Sentence Embedding task were categorized according to the four categories discussed in Section 2.4. Table 3 presents the produced responses per FB group on this task.

Table 3. Distribution of responses per FB group (FB2-passers, FB1-passers and FB-failers) in mean percentages (standard error) in the Sentence Embedding task

Response type	FB2-passers	FB1-passers	FB-failers	Total
Double embedding	57.8 (10.5)	20.0 (5.3)	17.2 (15.4)	29.8 (5.10
Mixed embedding	23.7 (8.5)	13.6 (4.2)	5.7 (5.7)	14.9 (3.4)
Single embedding	18.2 (8.4)	52.8 (7.1)	29.7 (16.4)	42.7 (5.7)
No embedding	-	11.0 (4.2)	47.4 (21.2)	12.6 (4.0)

Overall, single embeddings were produced most frequently (43%), followed by the targeted double embeddings (30%). There was a differential pattern across the three groups: the FB2-passers mainly produced double embeddings (58%), the FB1-passers mainly produced single embeddings (53%), and the FB-failers mainly produced non-embeddings (47%). Despite this differential pattern, most groups produced all sentence types, although the FB2-passers did not produce sentences without embedding. Notably, one participant in the group of FB-failers produced double embeddings for 94% of the target items.

A one-way ANOVA analysis was carried out with FB group as independent variable and the mean production of the different response types as dependent variable. Children with different levels of FB reasoning performed differently in the production of double embeddings ($F(2, 54)=5.28$, $p<.01$, $\eta^2=0.169$). Post-hoc analyses revealed that the FB2-passers produced significantly more double embeddings than the FB1-passers ($p<.01$). The three groups did not differ in their production of mixed embeddings, but there were differences in the production of single embeddings ($F(2, 54)=6.17$, $p<.004$, $\eta^2=0.192$). The

FB1-passers produced significantly more single embeddings than the FB2-passers ($p<.01$). The groups also differed in their production of sentences without embeddings ($F(2, 54)=6.44$, $p<.01$, $\eta^2=0.198$). The FB-failers produced more sentences without embeddings than children who mastered first-order FB ($p<.001$), and also more in comparison to children who mastered second-order FB ($p<.001$). We did not find differences between the four combinations of the verbs *think* and *say*.

3.3 *Regression analysis*

In the previous section, we looked at children's production of embedded sentences in the Sentence Embedding task. In this section, we consider how children's production of embedded sentences can be predicted by their scores on FB reasoning, reading ability and age.

First, a linear regression analysis was performed with the production of *double or mixed embeddings* as dependent variable and with FB1-scores, FB2-scores, DMT/AVI and age as explanatory variables. We combined double and mixed embedding because they both express recursive propositions (e.g., beliefs about beliefs). Furthermore, due to the collinearity between the two reading ability scores (DMT and AVI; $r(53)=0.87$, $p<.001$), one combined measure for reading ability was used (DMT/AVI). The model explained 40.2% of variance in the production of double/mixed embeddings ($F(4,54)=8.39$, $p<0.001$).[ii] Table 4 presents the standardized coefficients, showing that double/mixed embeddings are predicted by children's performance on the second-order FB questions $\beta=0.52$, $t(54)=4.23$, $p<.001$. The effect of reading ability failed to reach significance. FB1-score and age were no significant predictors for the production of double/mixed embeddings either.

Table 4. Linear regression model for predicting double/mixed embeddings from FB1-score,

FB2-score, AVI/DMT and age

Predictor	Beta	T	Sig.
FB1-score	0.015	0.136	0.893
FB2-score	0.52**	4.23	0.000
DMT/AVI	0.217	1.83	0.073
Age (years)	0.075	0.595	0.555

Note. **p<.001

Second, a linear regression analysis was performed with the production of *single embeddings* as dependent variable and with the FB1-scores, FB2-scores, DMT/AVI and age as explanatory variables. The model explained 32.1% of variance in the production of single embeddings ($F(4,54)=5.913$, $p<0.001$). Table 5 presents the standardized coefficients, showing that the proportion of correct FB1-answers failed to significantly predict the proportion of single embeddings, but that both a better reading level ($\beta=-0.35$, $t(54)=-2.74$, $p<.01$) and a higher rate of correct FB2-answers ($\beta=-0.35$, $t(54)=-2.71$, $p<.01$) were equally predictive of a lower rate of single embeddings. Age was no significant predictor for the production of single embeddings.

Table 5. Linear regression model for predicting single embeddings from FB1-score, FB2-score, AVI/DMT and age

Predictor	Beta	T	Sig.
FB1-score	0.22	1.80	0.078
FB2-score	-0.35*	-2.71	0.009

DMT/AVI	-0.35*	-2.74	0.008
Age (years)	-0.029	-0.22	0.827

Note. *p<.01

4. ***Discussion***

This study aimed to investigate the relation between children's production of double sentence embedding and second-order FB reasoning. The hypothesis tested is an extension of the language-first hypothesis by De Villiers and colleagues (e.g., De Villiers and Pyers, 2002), which holds that the acquisition of the syntax of embedded complements under mental-state verbs is a prerequisite for the development of first-order ToM. Extending this hypothesis, we wanted to know whether the ability to produce double embeddings is related to, and perhaps even required for, the mastery of second-order FB reasoning.

Our study shows first of all that second-order FB reasoning is still under construction in middle childhood. The children in our study (aged 7 to 10) performed at ceiling on the first-order FB questions, but still experienced difficulties with second-order FB reasoning.

In our elicited production task, children produced a considerable number of single and double embeddings. Their production of double embeddings was found to be related to second-order FB reasoning in particular. First, children's production of double embeddings was explained by their second-order FB performance, but not by their first-order FB performance, reading abilities or age. Second, children with different FB abilities showed different patterns of production of sentence embedding; only the children who mastered second-order FB understanding had a preference for producing double embeddings.

These results indicate that the ability to produce double embeddings and the ability of second-order FB reasoning are related in children. However, our study does not provide evidence that the ability to produce double embeddings is a prerequisite for second-order FB reasoning. Rather, the relation between double embeddings and complex FB reasoning in our study is not unique and exclusive. Double embeddings were produced by children at all levels of FB mastery, also by children who failed on the FB1 and FB2 questions. Furthermore, children who mastered second-order FB understanding did not mainly produce double embeddings; rather, they did so only 58% of the time and also produced a substantial number of mixed and single embeddings. As the Sentence Embedding task was feasible even for the youngest children, the children who produced single embeddings in this task likely had difficulty producing double embeddings. If so, this would suggest that the ability to produce double embeddings is not a prerequisite for second-order FB reasoning and hence that our results go against the language-first hypothesis for recursive ToM.

Our production task also revealed an unexpected strategy in children. Some children produced mixed embeddings, that preserved the complexity of mental-state embedding while avoiding the syntactic complexity of double sentence embedding. These mixed embeddings contain direct speech reports, which require less cognitive effort to produce than indirect speech reports (Groenewold, Bastiaanse, Nickels, Wieling and Huiskes 2014). This strategy of producing simpler mixed sentence embeddings for representing complex mental-state embeddings could be taken to indicate that the children are capable of second-order ToM, but still have difficulty representing this complexity in language by means of sentential complementation, which would again go against the language-first hypothesis for recursive ToM. However, perhaps the mental-state embeddings represented in the visual stimuli available to the child need not be represented mentally by the child. It is conceivable that the children producing mixed embeddings did so by step-wise verbalizing the mental states visible in the picture. Nevertheless, these mixed embeddings seem typical for children and may be a *rite de passage* towards producing full-blown double embeddings. However, confirmation of this suggestion requires further study.

Overall, the children in our study produced a large number of single embeddings. Many of these were produced by children capable of first-order but not second-order FB reasoning. This confirms the hypothesized relation between the mastery of single sentence embedding and first-order ToM. The finding that mastery of first-order FB understanding did not predict single embedding could be due to the children's ceiling level performance on first-order FB. An alternative explanation for the large number of single embeddings could be that the children were trained on producing single embeddings during the practice session. However, all children remarked at the start of the test session that the pictures had changed, and this was also explicitly mentioned by the experimenter. It is therefore unlikely that the children did not produce double embeddings because they failed to see that the pictures had changed.

This study is a first step into investigating the relation between complex language and complex thought. It reveals that there is a relation between the production of double embeddings and mastery of second-order FB reasoning. Future studies, investigating the longitudinal development of language and Theory of Mind, should uncover the exact direction of this relation.

5. References

Astington, Janet W. & Jodie A. Baird. 2005. *Why language matters for a Theory of Mind.* New York: Oxford University Press.

De Mulder, Hannah N. M. 2011. *Putting the pieces together: The development of Theory of Mind and mental language*. PhD diss., Utrecht University.

De Villiers, Jill G. 2005. “Can language acquisition give children a point of view?” *Why language matters for a Theory of Mind* ed. by J. Astington and J. Baird, 188-219. New York: Oxford University Press.

De Villiers, Peter A. 2005. “The role of language in Theory-of-Mind-development: What deaf children tell us”. *Why language matters for a Theory of Mind* ed. by J. Astington and J. Baird, 266-297. New York: Oxford University Press.

De Villiers, Jill G. & Jennie E. Pyers. 2002. “Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief understanding”. *Cognitive Development* 17:1.1037-1060.

De Villiers, Jill G., Kathryn Hobbs, & Bart Hollebrandse. 2014. “Recursive complements and propositional attitudes”. *Recursion: Cognitive Complexity* ed. by T. Roeper and P. Speas, 221-242. Vienna: Springer Verlag.

Groenewold, Rimke, Roelien Bastiaanse, Lyndsey Nickels, Martijn Wieling & Mike Huiskes. 2014. “The effects of direct and indirect speech on discourse comprehension in Dutch listeners with and without aphasia”. *Aphasiology* 28:7.862-884.

Hollebrandse, Bart, Angeliek van Hout & Petra Hendriks. 2014. “Children’s first and second-order false-belief reasoning in a verbal and a low-verbal task”. *Synthese* 191:3.321–333.

Hollebrandse, Bart & Angeliek van Hout. 2015. “Comprehension and production of double-embedded structures: A kindergardenpath in long-distance dependencies in Dutch school-aged children?”. *Language Acquisition and Development: Proceedings of*

GALA 2013 ed. by C. Hamann & E. Ruigendijk, 190-204. Newcastle upon Tyne: Cambridge Scholars Publishing

Hollebrandse, Bart, Kathryn Hobbs, Jill de Villiers & Tom Roeper. 2008. “Second order embedding and second order false-belief”. *Language Acquisition and Development: Proceedings of GALA 2007* ed. by A. Gavarró & M. João Freitas, 268-278. Newcastle upon Tyne: Cambridge Scholars Publishing.

SLO. 2008. Beschrijving AVI-niveaus primair onderwijs, <http://www.slo.nl/primair/leergebieden/ned/taalsite/lexicon/00029/ > (1 April 2016).

Wellman, Henry M., David Cross & Julanne Watson. 2001. “Meta-analysis of Theory of Mind development: The truth about false-belief”. *Child Development* 72:3.655-684.

Wimmer, Heinz & Josef Perner. 1983. “Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception”. *Cognition* 13:1.103-128.

---

[i] We report on the results of all 55 children, but find the same results if this one child is excluded from analysis.

[ii] A similar regression model with only double embeddings as the dependent variable explained slightly less variance (38%).