



# About the Relation Between Implicit Theory of Mind & the Comprehension of Complement Sentences

Dissertation

zur Erlangung des akademischen Grades  
Doktor der Philosophie (Dr. phil.)

eingereicht an der  
Humanwissenschaftlichen Fakultät  
Universität Potsdam  
Department Linguistik

von  
Silvana Poltrock

Potsdam, November 2010

This work is licensed under a Creative Commons License:  
Attribution - Noncommercial - Share Alike 3.0 Germany  
To view a copy of this license visit  
<http://creativecommons.org/licenses/by-nc-sa/3.0/de/>

Published online at the  
Institutional Repository of the University of Potsdam:  
URL <http://opus.kobv.de/ubp/volltexte/2011/5229/>  
URN <urn:nbn:de:kobv:517-opus-52293>  
<http://nbn-resolving.org/urn:nbn:de:kobv:517-opus-52293>

## Acknowledgments

This dissertation would not have been possible without the support of many people to whom I would like to express my sincere gratitude.

First of all, I am very grateful to Prof. Barbara Höhle for accepting me as her doctoral student, for introducing me to the interesting field of language acquisition and for her great support in all kinds of questions related to the preparation of the thesis.

I thank Prof. Birgit Elsner for supervising me and for her insightful comments from the developmental psychology perspective. Further, I would like to express my gratitude to Dr. Uli Sauerland who inspired me to conduct the second study of this thesis.

A special thank goes to my friend and colleague Frauke Berger for her support in the preparation of the experiments, for intense discussions on the findings and especially for motivating me during the final phase of the thesis.

I thank Tom Fritzsche for proof reading and for sharing his comprehensive knowledge about statistics and linguistics with me. Many thanks to Oda-Christina Brandt-Kobele and Gunnar Nehrke, who also made helpful comments on earlier versions of the thesis. Thanks for being continuously interested in my research and for encouraging me to proceed!

Further, I thank the student assistants who helped me to conduct the experiments: Agnes Groba, Katja Hummel, Caroline Magister, Maja Stegenwallner and Jana Weißbach.

For the financial support, I am grateful to the University of Potsdam whose funding enabled me to finish this thesis as a member of the *Graduate Programme for Experimental and Clinical Linguistics*.

Finally, I thank all the children who participated in the experiments and made it possible to explore these questions.





## Abstract

Previous studies on the relation between language and social cognition have shown that children's mastery of embedded sentential complements plays a causal role for the development of a *Theory of Mind* (ToM). Children start to succeed on complementation tasks in which they are required to report the content of an embedded clause in the second half of the fourth year. Traditional ToM tasks test the child's ability to predict that a person who is holding a false belief (FB) about a situation will act "falsely". In these task, children do not represent FBs until the age of 4 years. According the linguistic determinism hypothesis, only the unique syntax of complement sentences provides the format for representing FBs. However, experiments measuring children's looking behavior instead of their explicit predictions provided evidence that already 2-year olds possess an *implicit* ToM. This dissertation examined the question of whether there is an interrelation also between implicit ToM and the comprehension of complement sentences in typically developing German preschoolers.

Two studies were conducted. In a correlational study (*Study 1*), 3-year-old children's performance on a traditional (explicit) FB task, on an implicit FB task and on language tasks measuring children's comprehension of tensed sentential complements were collected and tested for their interdependence. Eye-tracking methodology was used to assess implicit ToM by measuring participants' spontaneous anticipatory eye movements while they were watching FB movies. Two central findings emerged. First, predictive looking (implicit ToM) was not correlated with complement mastery, although both measures were associated with explicit FB task performance. This pattern of results suggests that explicit, but not implicit ToM is language dependent. Second, as a group, 3-year-olds did not display implicit FB understanding. That is, previous findings on a precocious reasoning ability could not be replicated. This indicates that the characteristics of predictive looking tasks play a role for the elicitation of implicit FB understanding as the current task was completely nonverbal and as complex as traditional FB tasks.

*Study 2* took a methodological approach by investigating whether children display an earlier comprehension of sentential complements when using the same means of measurement as used in experimental tasks tapping implicit ToM, namely anticipatory looking. Two experiments were conducted. 3-year-olds were confronted either with a complement sentence expressing the protagonist's FB (Exp. 1) or with a complex sentence expressing the protagonist's belief without giving any information about the truth/ falsity of the belief (Exp. 2). Afterwards, their expectations about the protagonist's future behavior were measured. Overall, implicit measures reveal no considerably earlier understanding of sentential complementation. Whereas 3-year-olds did not display a comprehension of complex sentences if these embedded a false proposition, children from 3;9 years on were proficient in processing complement

sentences if the truth value of the embedded proposition could not be evaluated. This pattern of results suggests that (1) the linguistic expression of a person's FB does not elicit implicit FB understanding and that (2) the assessment of the purely syntactic understanding of complement sentences is affected by competing reality information.

In conclusion, this dissertation found no evidence that the implicit ToM is related to the comprehension of sentential complementation. The findings suggest that implicit ToM might be based on nonlinguistic processes. Results are discussed in the light of recently proposed dual-process models that assume two cognitive mechanisms that account for different levels of ToM task performance.

## Zusammenfassung

Zahlreiche Studien legen nahe, dass der Erwerb finiter Komplementsatzstrukturen eine Voraussetzung für die Entwicklung einer *Theory of Mind* (ToM) ist. Kinder beginnen in circa der zweiten Hälfte ihres 4. Lebensjahres Komplementsatzverständnisaufgaben zu lösen, die eine Wiedergabe des Komplementsatzes erfordern. Die ToM wird klassischerweise als erworben bezeichnet, wenn ein Kind korrekt vorhersagen kann, dass eine Person, die eine falsche Überzeugung hat, entsprechend "falsch" handeln wird. Es ist bekannt, dass sich diese Fähigkeit bei Kindern mit circa 4 Jahren entwickelt. Gemäß der Hypothese des linguistischen Determinismus liefert die syntaktische Komplexität von Komplementsatzstrukturen das Format, um falsche Überzeugungen mental zu repräsentieren. In ToM Aufgaben, die erfassen, wo Kinder eine zukünftige Handlung einer fehlinformierten Person antizipieren, konnte allerdings bereits schon bei 2-Jährigen das Verstehen falscher Überzeugungen nachgewiesen werden. Diese frühe Fähigkeit wird auch implizite ToM genannt, da diese Art der Aufgabe keine bewusste Entscheidung verlangt. Die vorliegende Dissertation untersucht erstmalig im Deutschen, ob der Zusammenhang zwischen dem Verstehen von Komplementsatzstrukturen und ToM auch für diese implizite Fähigkeit gilt. Zwei Studien wurden durchgeführt.

In *Studie 1* wurden die Leistungen von 3-Jährigen in einer klassischen (expliziten) ToM Aufgabe, einer impliziten ToM Aufgabe und in Komplementsatzverständnisaufgaben erhoben und auf korrelative Zusammenhänge hin getestet. Dabei wurde mittels eines Eye-Trackers das antizipative Blickverhalten gemessen und somit auf das Vorhandensein einer impliziten ToM geschlossen. Die Leistungen in den Sprachaufgaben korrelierten nicht mit den Blickdaten, obwohl beide Maße mit den Leistungen in der expliziten ToM Aufgabe in Beziehung standen. Unerwarteterweise konnte jedoch generell kein implizites Verstehen falscher Überzeugungen bei 3-jährigen nachgewiesen werden. Da die implizite ToM Aufgabe nichtsprachlich war, wird der Einfluss von Aufgabeneigenschaften auf die Elizitierung von impliziter ToM diskutiert.

*Studie 2* untersuchte, ob 3-Jährige ein früheres, implizites Verstehen von Komplementsatzstrukturen in Aufgaben zeigen, in denen antizipatorische Blicke anstelle von expliziten Antworten gemessen werden und damit das gleiche Maß verwendet wird wie in impliziten ToM Aufgaben. Zwei Experimente wurden durchgeführt. Der präsentierte Komplementsatz drückte entweder eine falsche Überzeugung des Protagonisten aus (Exp. 1) oder eine Überzeugung, deren Wahrheitsgehalt nicht bestimmt wurde (Exp. 2). Während bei 3-Jährigen kein Verstehen von Komplementsatzstrukturen, die eine falsche Proposition enthielten, nachgewiesen werden konnte, zeigten Kinder ab einem Alter von 3;9 Jahren, dass sie den komplexen Satz verarbeiten, wenn dieser keine falsche Proposition beinhaltet. Dieses Ergebnismuster spricht dafür, dass (1) der sprachliche Ausdruck einer falschen Überzeugung die

implizite ToM nicht elizitieren kann und dass (2) das Erfassen des rein syntaktischen Verstehens durch zusätzliche Realitätsinformation beeinträchtigt wird.

Zusammenfassend konnte in der vorliegenden Arbeit kein Hinweis dafür gefunden werden, dass die implizite ToM in gleicher Weise wie die explizite ToM mit dem Verstehen von Komplementsatzstrukturen einhergeht. Die Ergebnisse legen nahe, dass vielmehr nonlinguistische Faktoren bei der Entwicklung einer impliziten ToM eine Rolle spielen könnten. Die Resultate werden mit Blick auf aktuelle Zwei-Prozess-Modelle diskutiert.

# Contents

List of Figures	xi
List of Tables	xiii
<b>1 Introduction</b>	<b>1</b>
<b>I THEORETICAL BACKGROUND</b>	<b>7</b>
<b>2 Theory of Mind</b>	<b>9</b>
2.1 Traditional ToM tasks tapping explicit ToM: Verbal predictions . . .	11
2.2 Indirect tasks tapping implicit ToM: Eye gaze . . . . .	16
2.2.1 Anticipatory Looking . . . . .	16
2.2.2 Violation of expectation . . . . .	20
2.3 Explaining the gap: Accounts of the development of ToM . . . . .	22
<b>3 Language &amp; ToM</b>	<b>25</b>
3.1 Linguistic Determinism Hypothesis: Syntax first . . . . .	26
3.1.1 Emperical Evidence . . . . .	28
3.1.2 Criticism and alternative accounts . . . . .	35
3.2 Implicit FB understanding & Language . . . . .	38
<b>4 The Acquisition of Sentential Complementation</b>	<b>41</b>
4.1 Properties of Complement Sentences and CTVs . . . . .	41
4.2 Production Studies . . . . .	46
4.3 Comprehension Studies . . . . .	47
<b>5 Research Questions and Overview of the Studies</b>	<b>53</b>

<b>II</b>	<b>EMPIRICAL INVESTIGATIONS</b>	<b>57</b>
<b>6</b>	<b>Study 1:</b>	
	<b>Implicit ToM and the Comprehension of Complement Sentences</b>	<b>59</b>
6.1	Introduction/ Aims . . . . .	59
6.2	Participants . . . . .	61
6.3	Material & Procedure for Single Tasks . . . . .	61
6.3.1	Nonverbal Implicit ToM Task: Unseen Displacement . . . . .	61
6.3.2	Explicit Verbal ToM Task: Unexpected Content . . . . .	65
6.3.3	Complementation Tasks . . . . .	65
6.3.3.1	Truth-Value-Judgment Task . . . . .	66
6.3.3.2	Sentence-Picture Matching Task . . . . .	68
6.3.3.3	Reporting-the-Complement Task . . . . .	69
6.3.4	General Language Measures . . . . .	71
6.3.4.1	Test for Reception of Grammar: TROG-D . . . . .	71
6.3.4.2	Active Vocabulary Test: AWST-R . . . . .	72
6.4	General Procedure and Apparatus . . . . .	72
6.5	Data Analyses . . . . .	73
6.6	Results . . . . .	74
6.6.1	ToM Measures . . . . .	74
6.6.1.1	Nonverbal Implicit FB Task . . . . .	74
6.6.1.2	Explicit FB Tasks . . . . .	80
6.6.1.3	Relation between explicit and implicit ToM . . . . .	82
6.6.2	Language Measures . . . . .	83
6.6.2.1	General Language Measures . . . . .	83
6.6.2.2	Complementation Tasks . . . . .	84
6.6.3	Relations between FB tasks and language tests . . . . .	88
6.6.3.1	Explicit ToM and language measures . . . . .	88
6.6.3.2	Implicit ToM and language measures . . . . .	93
6.7	Discussion . . . . .	95
<b>7</b>	<b>Study 2:</b>	
	<b>Implicit Understanding of Complement Sentences</b>	<b>103</b>
7.1	Introduction/ Aims . . . . .	103
7.2	Experiment 1: Expressing a False Belief . . . . .	104
7.2.1	Participants . . . . .	104
7.2.2	Material . . . . .	105
7.2.3	Apparatus and Procedure . . . . .	107
7.2.4	Data Reduction . . . . .	107
7.2.5	Results . . . . .	108
7.2.6	Discussion and Motivation of Experiment 2 . . . . .	110

---

7.3	Experiment 2: Expressing a Belief . . . . .	112
7.3.1	Participants . . . . .	112
7.3.2	Material . . . . .	113
7.3.2.1	Eye-tracking Task . . . . .	113
7.3.2.2	Negation Comprehension Tasks . . . . .	115
7.3.3	Procedure and Apparatus . . . . .	118
7.3.4	Data Reduction . . . . .	118
7.3.5	Results . . . . .	119
7.3.5.1	Eye-tracking Task . . . . .	119
7.3.5.2	Negation Comprehension Tasks . . . . .	122
7.3.6	Discussion . . . . .	123
7.4	General Discussion . . . . .	127
<b>8</b>	<b>General Conclusions</b>	<b>129</b>
8.1	Summary . . . . .	129
8.2	Methodological Considerations . . . . .	132
8.2.1	Implicit FB understanding: What elicits it? . . . . .	132
8.2.2	Impact of reality information in complementation tasks . . . . .	133
8.3	Is Implicit ToM independent of syntactic knowledge? . . . . .	135
	<b>Bibliography</b>	<b>139</b>
	<b>Appendix</b>	<b>157</b>
	<b>Erklärung</b>	<b>167</b>





# List of Figures

2.1	Picture presented in a ToM task (taken from Woolfe et al., 2002). . . . .	9
2.2	Procedure of Southgate et al. (2007), taken from Doherty (2009). . . . .	18
3.1	Example of an item from the memory-for-complements task devised by de Villiers and Pyers (2002). . . . .	29
4.1	Simplified syntactic tree of the complement sentence: <i>Peter glaubt, dass sie den Kuchen aß.</i> . . . . .	42
5.1	Overview of the present studies within the context of previous research.	54
6.1	Visual scene of the implicit nonverbal ToM task . . . . .	62
6.2	Selected frames from the events of the implicit ToM task . . . . .	64
6.3	Example of a picture presented in the truth-value-judgment task. . . . .	67
6.4	Example of an item presented in the sentence-picture matching task . . . . .	69
6.5	Example of an item presented in the reporting-the-complement task . . . . .	70
6.6	Areas of Interest in the implicit nonverbal ToM task. . . . .	74
6.7	Average looking proportion during the course of trials of the implicit nonverbal ToM task . . . . .	76
6.8	Average looking proportion to each AOI within 1600 ms . . . . .	78
6.9	Average looking proportion to each AOI within 1600 ms for Smarties passers and nonpassers separately. . . . .	82
6.10	Average looking proportion to each AOI within 1600 ms for complement tasks passers and nonpassers separately. . . . .	95
7.1	Selected frames from Exp. 1: Expressing a false belief . . . . .	106
7.2	Exp.1: Mean proportion of correct anticipations within 1000 ms . . . . .	109
7.3	Exp.1: Frequency distribution of individual averaged proportion of correct anticipations (older 3-year-old children only). . . . .	110
7.4	Selected frames from Exp. 2: Expressing a belief . . . . .	114
7.5	Comprehension of Negation: Example of an item presented in the sentence-picture matching task . . . . .	116

7.6	Exp.2: Mean proportion of correct anticipations in each condition and each age group. . . . .	121
7.7	Exp.2: Mean proportion of correct responses in each negation comprehension task and each age group. . . . .	122

# List of Tables

6.1	Number of trials with correct and incorrect first looks, split by age group and condition . . . . .	77
6.2	Performance on the different explicit FB tasks split by age group . . .	81
6.3	Performance on TROG-D and AWST-R split by age group (children only) . . . . .	84
6.4	Mean proportion of correct responses on the complement tasks split by age group (children only). . . . .	86
6.5	Partial correlations between screening tests controlling for age (3-year-old children only). . . . .	88
6.6	Bivariate correlations between explicit ToM measures and language tests (3-year-olds only). . . . .	89
6.7	Results of hierarchical logistic regression analyses for sentential complement understanding predicting explicit ToM measures, controlling for age (3-year-old children only). . . . .	91
6.8	Contingency table showing numbers of children passing and failing the Smarties <i>other</i> task and the reporting-the-complement task (false complements only) (3-year-old children only). . . . .	92
6.9	Contingency table showing numbers of children passing and failing the Smarties <i>other</i> task and the reporting-the-complement task (true complements only) (3-year-old children only). . . . .	93
6.10	Partial correlations between each implicit ToM metric and language tests controlling for age (3-year-old children only). . . . .	94
7.1	Exp. 2: Mean proportion of correct looks and one-sample t-tests scores for all conditions and age groups. . . . .	120



# Chapter 1

## Introduction

‘‘The limits of my language are the limits of my world.’’

(Wittgenstein, 1922)

Both language and possessing a Theory of Mind (henceforth ToM) are very sophisticated capabilities and posited to be uniquely human. ToM is defined as the ability to understand and reason about one’s own and others’ mental states (desires, beliefs, knowledge etc.) which enables us to explain and predict behaviour (e.g., Sodian and Thoermer, 2006). Until now and still controversial, there is no evidence that primates or other species can be credited with a ToM (for a review, see Call and Tomasello, 2008). Traditional ToM tasks test children’s ability to attribute a false belief (henceforth FB) to another person and predict the person’s "false" behavior based on this FB (Baron-Cohen et al., 1985; Gopnik and Astington, 1988; Perner et al., 1987; Wimmer and Perner, 1983). Concerning language, the ability to process complex syntactic structures has been identified as a core feature which distinguishes humans from primates (Fitch and Hauser, 2004). This finding is in line with the assumption made by Hauser et al. (2002) that the ability to embed a structure into another structure - called *recursion* - defines the language faculty as a uniquely human skill. The question arises to what extent recursion is the only ability we need to explain other cognitive skills that are unique to humans like the development of a ToM. Is extra-linguistic structural complexity in social cognition derived from language or as David Premack (one of the first who investigated ToM in chimpanzees) asked: “Is language the key to human intelligence?” (2004, p.318).

There is no doubt that language and social cognition are in general intimately connected and that they interact with each other (for a review concerning language and ToM development, see Milligan et al. (2007)). However, the nature of the relationship is, ontogenetically speaking, a widely debated issue. In order to communicate

successfully, some understanding between participants must obtain. For example, joint attention<sup>1</sup> is critical to word learning processes (e.g., Bloom, 2002; Tomasello, 1986), perspective-taking abilities are needed for the acquisition of pronouns (Ricard et al., 1999), and a sufficient ToM is required for the mastery of epistemic modals conveying different speaker's certainty (Papafragou, 2002). In all these cases, the cognitive development has been found to precede the mastery of linguistic structures.

The *linguistic determinism approach*, on the other hand, stresses the causal effect of language on social cognitive development. In a series of publications, Jill de Villiers and her colleagues propose that the mastery of the syntax of mental state and communication verbs namely the processing of embedded sentential complements is a prerequisite for ToM development (de Villiers and Pyers, 1997; de Villiers, 2000; de Villiers and de Villiers, 2000; de Villiers and Pyers, 2002; de Villiers and de Villiers, 2003; de Villiers, 2005a, 2007). According to this hypothesis, this linguistic structure provides a representational format to represent falsity (e.g., “Benny thinks *there is a monster under his bed.*”). That is, the whole sentence remains true even the proposition in the embedded clause is false. Thus, the recursive syntax allows to think about other people's (false) beliefs. De Villiers considers recursion as fundamental to the emergence of social abilities: “In fact, this [author's note: the representation of possible worlds] may be the real utility, the functionality, of linguistic recursion” (2007, p.1869).

The linguistic determinism hypothesis has been supported in various kinds of experimental designs (correlational studies, training studies, longitudinal studies) and with different populations (autism, deaf subjects, SLI-children) (e.g., de Villiers and Pyers, 2002; de Villiers, 2005b; de Villiers et al., 2003; Hale and Tager-Flusberg, 2003; Lohmann and Tomasello, 2003; Schick et al., 2007; Tager-Flusberg and Joseph, 2005). Children start to pass standard ToM tasks at around the age of 4 years, slightly after acquiring the structure of sentential complement constructions.

However, the claim that ToM development can not proceed without this specific syntactic ability remains controversial, and both empirical and theoretical grounds for questioning this proposal have been put forth. One criticism affects the task tapping the comprehension of sentential complements. In their so-called *memory-for-complement task* (de Villiers and Pyers, 2002) children are asked to remember and report a false complement, like in: *She said she found a monster under her chair, but it was really the neighbor's dog. What did she say?* Adler (2002) and Ruffman et al. (2003) have argued that this task is not solely a language comprehension task, but relies upon the child's ToM ability, since they confront the child with a false belief or statement. Thus, they conclude, a correlation between the ability to report false complements and ToM task performance is in fact an artefact. Considering experimental

---

<sup>1</sup>Joint attention is defined as triadic interaction in which a child and an adult are attending to the same object (e.g., Tomasello, 1986).

evidence, other findings suggest that rather general syntax (Astington and Jenkins, 1999) or even broader, general language ability (semantics and syntax) and parental input about mental phenomena (Ruffman et al., 2002, 2003) assist the development of FB understanding. Furthermore, studies examining the linguistic determinism hypothesis in other languages<sup>2</sup> (e.g., German, Mandarin, Cantonese) show that it is not exclusively structural complexity but rather verb semantics that seem to have an effect on ToM development (Cheung et al., 2004; Perner et al., 2003; Tardif and Wellman, 2000; Tardif et al., 2004). Besides, Smith et al. (2003) could show that also the mastery of relative clauses correlates with FB understanding<sup>3</sup>. The authors argue that the mastery of two representations (i.e. the relative clause event is embedded into a matrix clause event) is crucial for solving the FB task. Thereby, the truth value of the representations does not play a role. Taking these findings together, it is a still unsolved question which aspect(s) of language exactly promote ToM development.

Studies examining the causal role of the comprehension of sentential complements in the development of children's understanding of FB have always used traditional ToM tasks that demand a conscious and thus explicit reaction from the children, i.e. a prediction of the protagonist's behavior. Some researchers, however, have recognized a distinction between an *explicit* and a preceding, *implicit* ToM (Clements and Perner, 1994; Low, 2010; Onishi and Baillargeon, 2005; Southgate et al., 2007). In studies that have investigated implicit ToM by using children's anticipatory looking as an indirect measure, already 2- and 3-year-olds anticipated the protagonist's upcoming behavior correctly but still gave incorrect verbal answers (Clements and Perner, 1994; Ruffman et al., 2001a; Low, 2010). Interestingly, Krachun et al. (2009) reported recently that apes showed some correct looking behavior in a nonverbal FB task pointing towards an implicit understanding of belief. Of course, this assumption needs further investigations. Yet, both the developmental pattern within human beings and the presumable presence of an implicit understanding in primates poses the question whether an implicit FB understanding is language independent or whether the linguistic determinism hypothesis can be extended to implicit ToM.

Ruffman (2004) suggested that implicit understanding of belief might be based on the (unconscious) detection of behavioral regularities whereas in tasks in which a conscious decision is required, language (both semantics and syntax) provides the means for reflecting on FBs explicitly. The primary aim of the current thesis is to investigate if children's comprehension of sentential complements is also linked

---

<sup>2</sup>In this context, cross-linguistic studies enable to separate purely syntactical effects from semantical effects and thus rule out language-specific phenomena.

<sup>3</sup>Relative clauses are structurally similar to complement clauses as both are embedded into a main clause, but (unlike complement clauses) relative clauses have always the same truth value as the main clause.

to their implicit FB understanding. This would suggest that a specific syntactic understanding provides the critical means for representing FBs in *any* kind of FB task.

Throughout the thesis, 3- to 4-year-olds have been tested as children in this age range have been found to be in a transitional developmental stage as most of them are not proficient in passing explicit ToM tasks yet but display an implicit FB understanding. The objectives of the present work can be subdivided into answering the following questions:

- i. Is there a correlation between implicit FB understanding and the comprehension of sentential complements?
- ii. Do children show an earlier comprehension of sentential complements, if methods are used similar to those that have been used to elicit implicit ToM, specifically eye-tracking?
- iii. Are syntactic properties of sentential complements alone sufficient for the relational link to FB understanding?

The first two points of interest concern the main aim of this thesis, namely to investigate the interrelation between the comprehension of sentential complements and implicit FB understanding. The first, correlational approach includes a simultaneous collection of children's performance on an implicit ToM task, on an explicit ToM task and on language tasks tapping the understanding of complement sentences (see next paragraph for further details about the language tasks). Eye-tracking methodology was used to assess implicit FB understanding by measuring participants' anticipatory eye movements while they were watching cartoons in which FB stories were enacted. The correlational study is reported as *Study 1*.

The second approach picks up the fact that an implicit ToM preceding the mastery of sentential complements challenges the linguistic determinism hypothesis. At least it raises the question if the language dependence hypothesis only holds for an explicit mentalistic understanding or if it can be extended to implicit ToM. One way to shed more light on this is to investigate children's comprehension of sentential complements with the same means of measurement as used in experimental tasks tapping implicit FB understanding: anticipatory looking. This allows to find out whether an earlier comprehension of complement sentences can be found with implicit measures. For that purposes, two experiments were conducted and reported as *Study 2*. In a first experiment, preschoolers' anticipatory looking behavior was recorded after they were presented with a sequence of two sentences: One, a complex sentence containing the protagonist's FB under the complement taking verb *glauben* ('think') and the other, a simple sentence describing the true state of affairs. In Experiment 2, children were presented only with a sentence expressing a person's



belief via a complement sentence without providing additional information about whether the belief matches with reality. The aim here was to identify the influence of a compelling knowledge about reality which is supposed to make the task harder irrespective of the syntactic complexity of the critical sentences. This influence is known in the ToM literature as *pull of the real* or *reality bias* (e.g., Mitchell, 1994; Birch and Bloom, 2003).

The issue raised under (iii) concerns the nature of the influence of language comprehension on ToM. Three language tasks were constructed to measure children's comprehension of complement sentences. No mental state verb like *think* or *know* etc. was used in the present work because the semantics of a mental state verb is likely to overlap with FB understanding itself. In order to avoid this confounding of syntactic and mentalistic knowledge, the perceptive verb *sehen* ('see') and the communication verb *sagen* ('say') were used in the complementation tasks. Another question in this regard is whether it is solely the recursive structure which enables us to think and reason about mental states or whether it is the potency to express false propositions in the complement clause. In contrast to previously used complementation tasks, two newly developed tasks (a truth-value-judgment task and a picture-sentence matching task) did not require the children to report a false proposition but rather to interpret the embedded clause as being the object of the matrix clause irrespective of the truth/falsity of the embedded proposition. Moreover, a modified version of the memory-for-complement task was created in which both true and false complements were included.

The present work is divided into two parts: a theoretical embedding of the research questions (chapter 2-5) and an empirical part presenting the two studies that have been conducted (chapter 6-8). Chapter 2 reviews the literature concerning children's ToM development. Methods used for the assessment of children's explicit and implicit FB understanding are described. Furthermore, approaches to explain the performance gap between both measures are introduced. Chapter 3 provides an overview of previous studies on the developmental link between language and ToM. The linguistic determinism hypothesis and its major criticisms are presented in detail. With regard to the relationship between implicit ToM and language, results of a recent study (Low, 2010) are highlighted. Chapter 4 outlines the properties of complement sentences and complement-taking verbs and gives an overview of studies that investigated the acquisition of sentential complementation, both from a receptive and productive point of view. An outline of the present studies including the research questions and hypotheses is provided in Chapter 5. Chapter 6 presents the material and results of *Study 1* in which a correlational research design was employed. Chapter 7 presents two experiments on children's implicit comprehension of complement sentences. The results of the present work will be summarized and their implications will be discussed in Chapter 8.



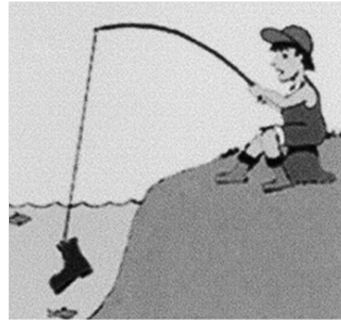
# Part I

## THEORETICAL BACKGROUND



## Chapter 2

# Theory of Mind



**Figure 2.1:** Picture presented in a ToM task (taken from Woolfe et al., 2002).

The term *Theory of Mind* was coined in 1978 by Premack and Woodruff who investigated whether chimpanzees possess the ability to attribute mental states to themselves and their conspecifics. The authors used the term *theory* because mental states are not directly observable but have to be inferred from observable behavior. A few years after Premack and Woodruff's seminal paper, developmental researchers extended this question to human development, i.e. asking when this specific social cognitive ability emerges during ontogeny.

ToM is an umbrella term including the understanding of multiple concepts of mental states (e.g., intentions, emotions, desires, knowledge) whose acquisition is assumed to follow a certain developmental progression (see the ToM scale developed by Wellman and Liu, 2004). In a more narrow definition, FB understanding is considered as the benchmark of the acquisition of a ToM. Wimmer and Perner (1983) devised a FB task in which the child is asked to predict an upcoming action of a

character who is holding a FB. The basis idea is that a child has an understanding of other's mental states (= a ToM) if s/he understands that a person can hold a belief that differs from reality and that causes this person's future incorrect behavior. In this thesis, children's FB understanding is used as the central measure of a mature ToM.

Children around the age of 4 years start to succeed on FB tasks that require verbal predictions (cf. a meta-analysis by Wellman et al., 2001). This rather late emergence is explained differently by various accounts. Several researchers claim that an important conceptual change takes place in the preschool years; before, the child has no concept of FB (e.g., Bartsch and Wellman, 1995; Gopnik and Wellman, 1992, 1994; Perner, 1991; Wellman, 1990). This account is also called the *Theory Theory* because children are assumed to infer behavior from beliefs analogous to a scientific theory. On the other hand, *early competence* accounts propose that children younger than 4 already have the ability to understand FB but the demands in standard FB tasks mask their competence as they require abilities other than FB understanding (e.g., Bloom and German, 2000; German and Leslie, 2000; Leslie, 1987, 1994; Leslie et al., 2005; Lewis and Osborne, 1990; Scholl and Leslie, 2001).

Studies using looking behavior as an indirect measure of FB understanding report evidence that children younger than 4 years are able to correctly anticipate the behavior of a person holding a FB (Clements and Perner, 1994; Low, 2010; Ruffman et al., 2001a; Southgate et al., 2007). Clements and Perner (1994) suggested that children's correct visual orienting reflects an *implicit* understanding of FB because a predictive looking task does not require a conscious decision from the child. Notably, these findings feed into the theoretical debate as the discrepancy between verbal responses and looking is consistent with *early competence* accounts. However, it is still under debate whether different means of measurement tap into the same FB reasoning ability (for discussion, see Perner et al., 2007) or whether the underlying mechanisms of children's early success in implicit FB tasks are the same as in explicit FB task, respectively.

As for explicit ToM, numerous studies demonstrated that children's acquisition of a specific syntactic construction, namely sentential complementation, is a precursor of ToM development (see Chapter 3). The present work aims to explore if children's implicit ToM is also linked to complement mastery which could, in turn, be an indication that both types of FB reasoning rest on the same foundation.

This chapter on ToM is subdivided into three sections. Section 1 reviews research on children's performance in FB tasks in which the child is required to make a verbal prediction. Performances on these tasks are defined to reflect explicit ToM<sup>1</sup>. Moreover, the impact of different task modifications (presentation mode, type of

---

<sup>1</sup>In the current work, tasks tapping explicit ToM are also referred to as explicit or direct FB tasks.

response) is presented. In the second section, FB tasks using children's eye gaze as a spontaneous response are described in detail. This type of task is assumed to tap into precocious knowledge referred to as implicit ToM<sup>2</sup>. There are two sources of evidence: First, children's performances in *anticipatory looking* tasks are presented (2.2.1) and second, results of studies employing the *violation-of-expectation* paradigm are reviewed (2.2.2). Finally, different approaches that aim to explain the performance gap between implicit and explicit tasks are discussed.

## 2.1 Traditional ToM tasks tapping explicit ToM: Verbal predictions

There are two widely used FB tasks which both require the child to make a (conscious) prediction about what a story's character will do next based on the character's misrepresentation of (1) an object's location (unseen-displacement task/ change-of-location task) or (2) the content of a container (unexpected-content task/ deceptive box test).

In the *unseen-displacement task*, also known as *Maxi task* (Wimmer and Perner, 1983) and *Sally-Anne task* (Baron-Cohen et al., 1985), children watch a scene in which the story's protagonist (e.g., Maxi) places an object in location A (e.g., a green cupboard) and walks away. While the protagonist is absent another character (e.g., Maxi's mother) takes the object and puts it in Location B (e.g., a blue cupboard) and leaves the scene. Then the story's protagonist (Maxi) comes back, and the experimenter asks the child participant: *Where will s/he look for the object?* If the child answers *in location A* (or points to location A), it passes the belief question by taking the protagonist's FB into account. If, however, the child answers *in location B* (or points to location B), then it fails the question by not appreciating the mental state of the protagonist.

In the second task, the *unexpected-content task* which is also known as the *Smarties task* (Gopnik and Astington, 1988; Perner et al., 1987), a Smarties tube is shown to the child who is then asked what s/he thinks is in the box. After the child answers usually *Smarties* or *sweets* the experimenter opens the box and shows that the tube in fact contains pencils. The experimenter closes the box again and asks the child (a) what s/he was thinking was inside the box before it was opened (*self* FB question) and (b) what someone else who had not yet seen the tube's content will think is inside the box (*other* FB question). A child passes the task if s/he answers *Smarties* or something comparable indicating that s/he can ascribe a (former) FB to oneself and/ or to another person. Instead, children demonstrate a lack of FB understanding when they respond with the real contents of the box: *pencils*.

---

<sup>2</sup>These tasks are also denoted as implicit or indirect FB tasks in the following.

In their meta-analysis in which 178 studies are included, Wellman et al. (2001) reported that children younger than 3 years and 6 months were performing consistently below chance, whereas children aged about 4 years were performing above chance (75% of the children were correct at 4 years and 8 months). The type of task, nature of the protagonist (e.g., a doll, videotaped or real character), nature of target object and type of question did not modify performances on FB tasks.

Theoretically motivated, many studies have been focusing on the question of whether other experimental modifications improve especially younger children's FB task performance. Theorists who assume a conceptual competence change during the preschool years (Theory Theory) hold that FB understanding is independent of the surface features of the task (e.g., Bartsch and Wellman, 1995; Gopnik and Wellman, 1992, 1994; Perner, 1991; Wellman, 1990). In contrast, theorists who assume rather an innate, early ToM competence which is masked by performance factors in traditional FB tasks (ToM module theory) predict considerably earlier FB understanding in tasks with reduced language and executive demands (e.g., Bloom and German, 2000; German and Leslie, 2000; Leslie, 1987, 1994; Lewis and Osborne, 1990; Scholl and Leslie, 2001).

### **Impact of task characteristics**

Wellman et al. (2001) identified five experimental factors that have an impact on children's FB task performance. Children performed better (1) if the time frame was emphasized (e.g. *Where will s/he look first for the object?*) - however, this effect was visible only for older children; (2) if deception motivated the change of the object's location (i.e., when the hider acted sneakily); (3) if children carried out the transformation themselves; (4) if the target object was not present when the FB question was asked and (5) if the protagonist's belief was stated explicitly. Nonetheless, none of these five experimental factors reliably raised young children's FB performance above chance level. Brown and Bull (2007) reported that 3-year-olds show no above-chance performance even if all influencing variables (1-4) were employed in the same FB task. Wellman et al. concluded that their results are compatible with the *conceptual change* account rather than with *early competence* account.

In their commentary on Wellman et al. (2001), Scholl and Leslie (2001) argue that these data do not speak against an *early competence* account but, on the contrary, support it as several task characteristics boost children's performance. Similarly, Moses (2001) questions whether the results of the meta-analysis really rules out the possibility that young children may have a concept of mental states but are unable to express that since "certain types of studies were entirely excluded from the [...] meta-analysis." (p.688). For example, Wellman et al. (2001) did not include studies using participants' own deceptive behavior as a measure of FB understanding (e.g.,



Chandler et al., 1989) or studies including FB explanation tasks (e.g., Wellman and Bartsch, 1988). Most important for the present work, all studies included were first of all language demanding, that is children were required to follow a narrative and to comprehend the verbal test question. Second, the meta-analysis does not include the feature *type of response*. Only studies that measured children's explicit judgments about an upcoming behavior were analyzed; studies using looking time measures or behavioral measures as an indicator of FB understanding were not considered. The impact of both task characteristics are considered in the following.

**Linguistic nature of FB task** Several researchers have developed FB tasks that require no or less linguistic resources in order to disentangle the relation between ToM and language. There are at least two possible ways in which language might affect FB task performance: linguistic demands in verbal-based FB tasks could mask children's FB competence and thus underestimate children's performance (Call and Tomasello, 1999; Chandler et al., 1989). On the other hand - as Plaut and Karmiloff-Smith (1993) proposed - language-free tasks might be harder because they lack verbal description which provides a scaffolding for representing FB situations. So far, there is no clear evidence in the literature concerning the question whether nonverbal ToM tasks are less or more difficult than verbal ones (for a review, see Astington and Baird, 2005a).

There are three parts of a FB task that can be nonverbal: the presentation, the question and the response. Call and Tomasello (1999) found no remarkable earlier onset of the mastery of a completely nonverbal hide-and-peek FB task compared to a less-verbal task where the test question was posed verbally but still the presentation of the FB situation was nonverbal. Thus, no effect of the linguistic nature of the test question was observed. However, in a study by Lunn (2003, reported in Astington and Baird, 2005a) in which both Call and Tomasello's (1999) tasks and additionally a traditional FB task were used, 3-year-olds performed better in both versions of the nonverbal/ less-verbal tasks compared to the standard FB task (app. 70% correct responses in the Call & Tomasello tasks vs 40% correct responses in standard FB task).

Astington and Baird (2005a) investigated the role of the presentation mode in a FB task. They presented 3-year-olds with three versions of the unseen-displacement-task in a within-subject design (i.e. each participant received all conditions): Besides the standard version in which the object displacement was shown and the story was narrated by a voiceover, children saw a *verbal-only* version (the object transfer was only narrated, but not seen) and a *visual-only* version (the object transfer was shown without any verbal input). Children then were asked to predict where the protagonist will look for the object, that is the question and the response were still verbally-based and thus explicit. There were no differences between the three conditions suggesting

that verbal input or its absence per se does not influence FB task performance. In contrast, Norris and Millan (1991, cited by Plaut and Karmiloff-Smith, 1993) reported that children performed worse (compared to a standard version) in a task in which the story was presented only visually without verbal input. Furthermore, there is evidence that presenting the FB story only verbally (i.e. narrating the unseen displacement story) makes the task easier compared to the standard version (Johnson and Maratsos, 1977; Zaitchik, 1991). These results indicate that narratives help children to represent the FB, possibly by guiding the child's attention to the critical aspects of the FB story. Moreover, the absence of visual information might enhance children's performances because the object's actual location is less compelling in this case. In conclusion, the question of whether an additional verbal description enhances children's FB task performances is not answered clearly in the literature.

Hahn (2009) presented 3-, 4- and 5-year-old children with both a verbal FB task and a completely nonverbal FB task in which the child was encouraged to act out the end of a FB story. Although a general similar developmental trajectory between verbal and nonverbal FB tasks was found, that is 4- and 5-year olds but not 3-years olds performed above chance in both FB tasks, 3-year-old children performed significantly better on the nonverbal FB task compared to the verbal FB task. Hahn (2009) concluded that the linguistic nature of ToM tasks additionally makes it more difficult for 3-year-olds to pass traditional FB tasks. Notably, there are at least two factors which could have had an impact on children's performance in the nonverbal version: The absence of verbal input and the fact that the child is required to act out the character's search instead of predict it verbally. This leads to the second task characteristic that has not been accounted for in the meta-analysis: type of response.

**Type of response** The effects concerning the type of response of a FB task are much more consistent. Results of studies using dependent variables other than verbal predictions, e.g. behavioral reactions or looking measures, show evidence that children younger than 4 years of age are sensitive to belief states of other individuals (e.g., Carpenter et al., 2002; Clements and Perner, 1994; Freeman et al., 1991; Happé and Loth, 2002; Onishi and Baillargeon, 2005; Southgate et al., 2007). 3-year-old children for instance were able to act out the protagonist's search pattern but could not give the correct prediction (Freeman et al., 1991).

Success in 3-year-olds has also been observed in FB tasks in which the child is required to identify the correct referent in a word-learning task (Carpenter et al., 2002; Happé and Loth, 2002). In these studies, the child, together with an experimenter E1, watches a novel object being hidden in one of two locations. As in the traditional unseen-displacement task, the object is transferred afterwards by another experimenter E2 while E1 is absent. A second novel object is placed in the first location. After E1 returned to the scene s/he wants to retrieve the

(first) object and labels it. The task for the child is to identify the object E1 is referring to. Children aged between 2;8 and 3;3 years performed significantly better compared to the baseline expectancy, i.e. the performance in a true belief (TB) condition (Carpenter et al., 2002). In the TB condition, which is often used as a control condition in FB tasks, the displacement of the object is seen also by the protagonist. Hence, the real state of affairs and the content of the protagonist's belief are identical.

Happé and Loth (2002) state that tracking the FB of E1 in a word-learning task is much easier than linking a mental state and a future behavior when making action prediction. In this kind of tasks, children are still required to make a decision but this decision is rather behavioral and indirectly measured. Happé and Loth (2002) themselves explain the earlier successful performance with the "implicit nature of the question regarding the character's belief" (p.31). Thus, a communicative, more natural context might be more suited to test for FB understanding. Still, in both studies (Carpenter et al., 2002; Happé and Loth, 2002), there is no evidence that children younger than 3 years of age show above-chance performance. In a recent study by Southgate et al. (in press) however, already 17-months old infants were able to assign reference correctly when a character is holding a FB about an object s/he is referring to. The procedure was basically the same as in previous implicit reference tasks but crucially the deceptive nature of the situation was emphasized, i.e. E2 behaved in a sneaky manner. As mentioned above, this task characteristic was found to increase children's performances also in standard FB tasks (Wellman et al., 2001).

A second body of research focusing on children's early FB competence has been using the helping-paradigm (Buttelmann et al., 2009; Matsui and Miura, 2008). Buttelmann et al. (2009) measured how 2;5-year olds, 16- and 18-month-old children helped either a misinformed (FB) or an informed (TB) experimenter to achieve his goal. Both 18-month- and 2;5-year olds behaved differently in each condition, that is they were able to take the experimenter's belief into account when inferring his goal and thus showed FB understanding. 16-month-old infants' responses were similar, but failed to reach significance. Further evidence for early ToM in prosocial actions were also found in 3-year-olds (Matsui and Miura, 2008).

A third line of research showing early FB understanding stem from indirect tests that measure at which age children start to *look* appropriately. Employing these tasks, a sensitivity of other's FB can be reliably demonstrated before the age of 3 years (Clements and Perner, 1994; Onishi and Baillargeon, 2005; Southgate et al., 2007). The following section gives an overview about studies using eye gaze to determine children's *implicit* ToM.

## 2.2 Indirect tasks tapping implicit ToM: Eye gaze

In general, two kinds of indirect FB tasks have been employed using children's eye gaze as a spontaneous response to reveal implicit ToM: *anticipatory looking* (AL) tasks (also called predictive looking tasks) and *violation-of-expectation* (VOE) tasks. In both paradigms, children's ability to understand other's mental states is inferred from their looking behavior when watching a FB scenario. As Ruffman (2000) pointed out, children's spontaneous response as indexed by gaze direction is supposed to reflect a social sensitivity in real social situations. Hence, Ruffman (2000) argues, focusing on behavioural implications of FBs is the more insightful way of testing children's FB understanding compared to eliciting a conscious prediction.

### 2.2.1 Anticipatory Looking

The ability to anticipate future actions seems to be present quite early in life: Infants as young as 4 months for instance are able to anticipate the reappearance of an occluded object (e.g., Johnson et al., 2003; Rosander and von Hofsten, 2004). Clements and Perner (1994) were the first ones who analyzed children's anticipatory looking behavior in a FB task as an index of unconscious and (as the authors labeled it) *implicit* ToM. They accomplished the standard unseen-displacement task and evoked anticipatory eye movements with a verbal prompt question (*I wonder where he is going to look?*) before they asked the child to predict where the story's protagonist (e.g., a mouse) will look for the object. Basically, the scenarios included two exits (e.g., left and right mouse hole), each with a corresponding box in front of it. The participants were videotaped while the story was narrated and enacted on cardboards. This recording was used later to determine where the child looked at immediately after the prompt was given. After the prompt question there was a pause of 1 to 2 seconds in which the participants had the chance to anticipate the reappearance of the protagonist at one of the two locations. A TB condition was included to rule out other interpretations of the eye movement pattern (e.g., retracing the sequence of story events).

Children from the age of 2;11 years on showed already anticipating gazes towards the correct (i.e. belief-accordant) location in both conditions but still gave incorrect verbal answers in the FB condition. Clements and Perner (1994) suggested a different type of knowledge which is unverbalizable and unconscious - two criteria of implicitness (Dienes and Perner, 1999). The results are in line with a general assumption about cognitive development, that is, implicit knowledge precedes explicit knowledge (Karmiloff-Smith, 1992).

The correct looking pattern could be replicated by Ruffman et al. (2001a) who used a 4-second period after the prompt question. In addition, they reported a decrease in certainty with increasing age for children who showed a mismatch between

looking measures (correct anticipation) and verbal reactions (incorrect prediction). The authors interpret this as providing evidence for a transitional stage in which children use different strategies. Furthermore, their data show that eye movements indeed index unconscious FB understanding as children are unaware of the knowledge conveyed by their anticipatory looks.

A study by Garnham and Ruffman (2001) provides evidence that children attribute a FB rather than simply ignorance to the agent. They presented three instead of two possible hiding locations and tested how specific the children's expectation concerning the incorrect future behavior was. Three-year-olds looked to that location in which the protagonist thought the object was hidden more often than to the third, irrelevant location. The authors conclude that an application of a *not seeing = not knowing rule* can be ruled out as an explanation for children's success in predictive looking tasks. That is, children do not simply expect the protagonist to search somewhere else but in the object containing box just because the protagonist did not see the transfer. Rather, they attributed a specific FB to her which leads to a certain behavior. Hence, these studies provide further support that anticipatory looking is a powerful measure of implicit FB understanding.

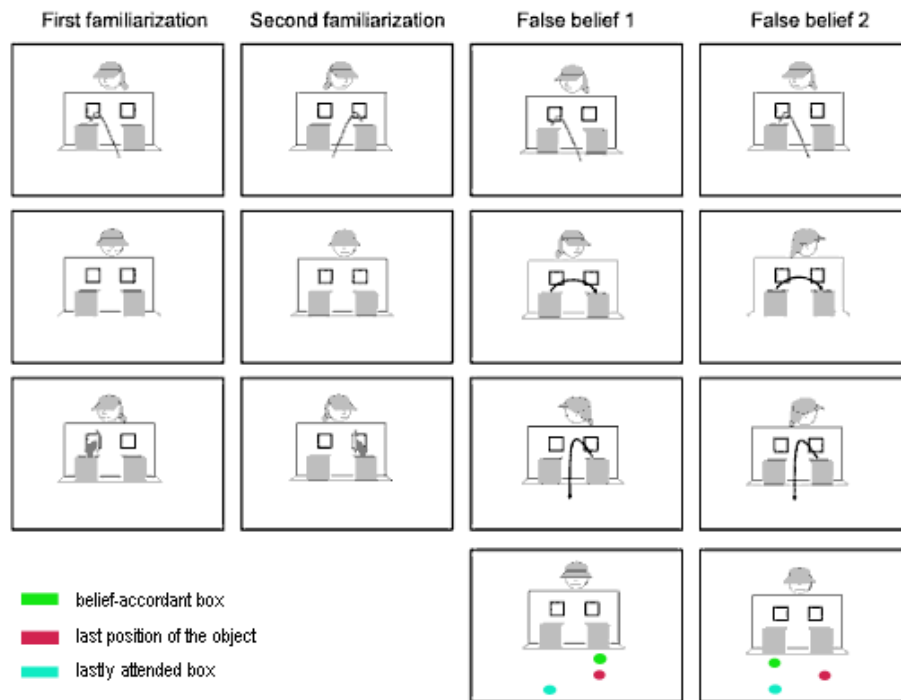
In a completely nonverbal AL task, Southgate et al. (2007) showed 25-months-old toddlers first two videos in which an actor watches a ball being hidden by a puppet in one of two boxes. Then the actor retrieves the ball by reaching through one window to that box containing the object (see first two columns of Figure 2.2). 1750 ms before the hand reaches to the box, both window frames lit up and a simultaneous sound occurred to elicit anticipatory eye movements. In a critical FB movie the actor again watches where the ball is hidden by the puppet, but then is distracted because of a ringing telephone while the puppet removes the ball from the scene. Then, the actor turns back and the sound and illumination occur to signalize the upcoming act of reaching. The removal of the object differs from the procedure of Clements and Perner (1994) and was done to eliminate the *pull of the real or reality bias* which describes children's problems with inhibiting the information of the real location of the object<sup>3</sup> (Birch and Bloom, 2003; Carpenter et al., 2002; Leslie et al., 2005).

To ensure that toddlers' gazes were not simply due to low-level cues, e.g. children just looked to the last position of the object, two FB movies were prepared and presented to two different groups of children: one group watched FB 1 (see Figure 2.2) in which the character is distracted after the second hiding, that is the last position of the ball is identical with the correct/ belief-accordant box (red circles in Fig. 2.2). In FB 2, however, the character turns around already after the first hiding, that is the last position of the ball is not identical with the correct/ belief-accordant box.

In addition, the two conditions differ with respect to the last position of the

---

<sup>3</sup>Hence, there was no TB condition included in Southgate et al., 2007.



**Figure 2.2:** Procedure of Southgate et al. (2007), taken from Doherty (2009).

character's attention (blue circles in Fig. 2.2) to ensure that children did not respond only to this low-level cue (i.e., looking to the box the protagonist lastly attended to). In FB 1, the actor lastly attended to the box where the toy was NOT placed in the end (i.e., the left one) before she turned away from the scene. In FB 2, she turns away after attending to the correct, belief-accordant box (also the left box).

Southgate et al. (2007) analyzed two measures of anticipation: the direction of the first saccade following the sound/illumination and the amount of time spent in each window<sup>4</sup>. In both FB conditions toddlers correctly anticipated the protagonist's action according to her FB as reflected in both dependent measures (for similar results with 18-month-old infants see Neumann et al., 2008). Interestingly, the authors avoided to label this ability *implicit* knowledge. This line of research strongly suggests that traditional FB tasks underestimate children's ability to forecast actions that are based on a character's false representation of the world.

However, Yi (2009) could not replicate these findings on early implicit FB reasoning. Since in the current work, a predictive looking task similar to the one designed by Yi

<sup>4</sup>Note that only children who anticipated correctly in the second familiarization trial were included in the analysis.

(2009, Exp.1) was used, the material and results are reviewed in more detail in the following. In experiment 1, he presented 3- and 4-year-olds and adults a nonverbal cartoon showing the unseen-displacement scenario but unlike Southgate et al. (2007) the object was not removed from the scene. The protagonist, a little girl, hides her toys either behind a rock or a lawn chair (the girl has to follow a Y-shaped path to reach either location). Then she walks back and disappears in the forest. While her eyes are covered the object is moving to the other hiding place by its own (FB trials) or is staying at the original place (TB trial). Then she turns back, starts to move and is occluded for 5 seconds behind the forest - the time period in which anticipatory eye movements were measured. To elicit an additional verbal answer participants were asked to predict where the protagonist will look for the toy at the end of one FB trial.

In the TB trial, 3- and 4-year-old children and adults correctly anticipated the protagonist to show up at the appropriate hiding place (i.e. where the toy still was). In the FB trials, 4-year-old children did not display a clear predictive looking pattern while 3-year-olds looked longer to the incorrect location (i.e. where the object really was) and adults looked more often towards the correct/ belief-accordant place. Further analysis showed that looking pattern in 4-year-olds were associated with their explicit prediction performances: children who answered the prediction question correctly also gazed more often towards the correct side; in turn children who failed to answer the explicit prediction question spent more time looking at the incorrect location.

Moreover, in experiment 2, an unexpected-content scenario was presented in four video clips of real people: a protagonist is asked about the content of a familiar looking box which he either opened (TB) or did not open (FB). Critically, the protagonist is hesitating while a close-up picture of the actual content (e.g. candles) and the expected content (crayons) is shown for five seconds. Following Yi (2009), eye movements during that period are assumed to index FB understanding since participants presumably look at those objects they think the story's character will point at. Again, results of the FB trials showed no clear looking preference for the 3- and 4-year olds but a late starting preference to the correct picture region for adults. Yi thus concludes that "the early understanding of false belief found in infants may be only applied to narrow and constrained situations." (p.79).

To summarize, AL experiments investigating implicit ToM differ with respect to their linguistic requirements (verbal vs. nonverbal) and in whether the object was removed from the scene or not before the anticipation phase starts. Perner et al. (2007) argue that it is, however, crucial to leave the target object in either of the two hiding places when assessing children's FB understanding because otherwise there is no *referential ambiguity*, that is a competition between the object's real location and the assumed/ last witnessed location. Basic processes then could subserve children's correct looking behavior instead of truly metarepresentational skills.

### 2.2.2 Violation of expectation

A second methodological approach to probe for infants' concept of belief is the violation-of-expectation paradigm (VOE). Basically, VOE studies test whether infants look considerably longer to impossible events as opposed to possible events. Thus, looking times after an event is interpreted as a measure of expectation: Infants tend to look longer to an event that contradicts what they would normally expect (Baillargeon et al., 1985).

The first attempt to assess infants' FB reasoning with a VOE task was made by Onishi and Baillargeon (2005). They presented 15-month-old infants movies in which an actor either had a FB or a TB and measured how long they looked after the actor either acted according to her belief (expected event) or contrary to her belief (unexpected event). At the beginning, each infant received three familiarization trials showing (1) the actor playing with a toy and putting it in one of two boxes; (2+3) the actor reaching inside that box for the object. Next, each infant received one trial that varied across condition: In a first FB condition, the actor does not witness that the toy moves from the right-hand box to the left-hand box on its own. In a second FB condition, the actor witnesses the first change of location, but is absent when the toy moves back to the original location. In the first TB condition, the actor witnesses that the object moved from one box to the other box; and in the second TB condition the box containing the object moves back and forth without a change of the object's location. Finally, in a test trial the actor reaches either into the left-side or the right-side box.

In all conditions, 15-month-olds looked reliably longer at displays in which an actor's search for a toy was inconsistent with her belief about the toy's location. This indicates that infants expected the actor to search for the object according to her belief about its location. Onishi and Baillargeon (2005) concluded that infants are able to track what an actor can perceive and thus knows and that they understand that someone's representation of a situation (both false and true) determinates how this person will act. Similarly, Surian et al. (2007) have shown that even 13-month-old infants are surprised (as conveyed in their looking times) when a (nonhuman) agent searches in the place where the object really is although the agent did not witness the transfer.

Perner and Ruffman (2005) and Ruffman and Perner (2005) question the conclusions of Onishi and Baillargeon (2005) and give an alternative interpretation of the results: Increased looking times can also stem from neurological activities or from low-level heuristics and do not necessarily demonstrate infants' understanding that the mind guides the behavior. Perner and Ruffman (2005) proposed that infants may form a three-way actor-object-location association during the belief-induction trial which causes shorter looking times when they encounter old combinations in the testing trial compared to new combinations. Alternatively, they claimed that



acting on a simple behavioral rule such as "someone always searches where s/he last saw the object" could also lead to correct expectations without a deeper FB understanding. However, Träuble et al. (2010) tested whether children act solely based on a seeing-searching rule. They presented an additional belief-induction context to 15-month-olds: an female actor could not see the object's displacement but only accessed this information manually by touching a balanced beam and causing the object's transfer to another box herself while facing a different direction. Participants looked significantly longer after a testing trial in which the actor reached for the empty box (incorrect location) compared to a trial in which the actor reached to the box containing the object (correct location). Träuble et al. (2010) argue that infants are able to track an agent's mental state in a flexible way which weakens the proposal of a "search-where-you-last-looked"-rule.

In addition, there are several recent VOE studies that demonstrate that infants' FB sensitivity is not restricted to special FB contexts but extends to different experimental situations involving mental state attribution (pretense: Onishi et al., 2007; object identity: Scott and Baillargeon, 2009; false perceptions: Song et al., 2008; unexpected content task: He et al., in press). Moreover, this capacity for tracking belief states seems rather flexible: Song et al. (2008) showed that 18-month-olds understand that an actor's FB about an object's location can be corrected by an appropriate communication.

In summary, this growing body of research using the VOE paradigm provides evidence that infants in their second year of life are able to discriminate possible events from events that are impossible because a person is acting contrary to her/his mental representation. Perner et al. (2007), however, calls into question if looking times really reflect expectations, and proposes instead that they might rather reflect processing costs. That would imply that these tasks might not tap ToM in terms of metarepresentational understanding but, again, in terms of superficial behavioral regularities or learning associations without inferring a mental state underlying the behavior. In order to account for the criticism concerning VOE experiments (see also Kagan, 2008, for a summery of critiques of the VOE paradigm), a predictive looking task was chosen in the current work to assess children's implicit ToM. That is, children's specific expectations (instead of changes in their looking times) are considered as providing the most compelling measure of implicit FB understanding.

However, the question is still unsolved why young children still fail on traditional, direct FB tasks although they are sensitive to other's FBs in indirect tasks. The next section gives an overview of explanatory models of the ToM development accounting for this performance gap between explicit and implicit FB tasks.

## 2.3 Explaining the gap: Accounts of the development of ToM

The findings on an early, implicit FB awareness in children evoke different reactions among researchers. On the one hand, traditionalists warn against a mentalistic interpretation of the data. They insist that children start to understand FB not until the age of around 4 years and that these precocious signs of FB understanding in infants can be explained by exploiting statistical regularities in sequences of behaviors (Perner et al., 2007; Perner and Ruffman, 2005; Ruffman and Perner, 2005). Perner et al. (2007) argue that also nonhuman animals show social competence but without establishing a mentalistic understanding of how the mind mediates between visual access and outcome (see also Povinelli and Vonk, 2003, 2004). This could apply to infants, too.

Other authors, however, claim that infants do have a concept of belief as reflected in their spontaneous looking behavior and that this implicit knowledge provides the conceptual foundation for children's metarepresentational ability that enables them to master direct FB questions (Onishi and Baillargeon, 2005; Leslie et al., 2005; He et al., in press; Baillargeon et al., 2010). According to their *response account* (Scott and Baillargeon, 2009; Baillargeon et al., 2010), there are two additional processes required in a task in which the child is asked to predict a behavior explicitly: Besides representing a FB, children also have to simultaneously (a) select a response and (b) inhibit the prepotent response (i.e. inhibit their own knowledge). Since in indirect tasks only the first, belief-representation process is involved, children remarkably younger than 4 years of age can succeed. Proceeding neurological maturation is assumed to cause the successful communication between these three processes (representing the FB, response-selection, inhibition of the prepotent response) and thus the mastery of predictive FB questions (He et al., in press; Baillargeon et al., 2010).

Apperly and Butterfill (2009) explain the performance gap by taking a different perspective and drawing an analogy with number cognition. There is evidence that reasoning about numbers includes two kinds of cognitive mechanism: One that emerges early in development, that is language-independent and accounts for limited<sup>5</sup> numerical competence in preverbal children and nonhuman animals, and a second one that develops later, is cognitively more demanding and highly language-dependent (see also Gallistel and Gelman, 1992, 2000; Gelman and Gallistel, 2004). The idea is that language provides the means which enable humans to represent numbers precisely and operate on them. Apperly and Butterfill (2009) propose that the existence of two systems also holds for FB reasoning: The early, nonverbal ToM

---

<sup>5</sup>The nonverbal capacity is limited to three to four items and rather prone to be unprecise (e.g., Le Corre and Carey, 2007).

mechanism works "quick and dirty", that is efficient but limited, and is reflected rather in children's communicative and interactive, spontaneous behavior. The second system heavily relies on higher cognitive functions like language and executive functions and thus demands more processing resources that might not develop until a certain age. Both systems exist assumedly in parallel in older children and adults.

The role of language in the origin of explicit FB understanding is also stressed by Ruffman (2000, 2004) and de Villiers and de Villiers (2003) who propose that implicit and explicit insights are based on different processes. De Villiers and de Villiers (2003) suggested that implicit FB understanding could be based on simulation or empathy - a system that human beings probably share with social primates - and that language comes into play when explicit reasoning is required. According to Ruffman (2004), implicit ToM builds on statistical learning processes whereas explicit ToM development is promoted by language as it provides the representational means to reflect on implicit insights (see also Ruffman et al., 2003). A training study by Clements et al. (2000) provides evidence for this assumption: Only children who exhibit implicit FB understanding could improve their explicit FB understanding after two training sessions<sup>6</sup>. Low (2010) showed recently that children's language skills, more precisely their mastery of sentential complementation, are correlated with their verbal responses in traditional FB tasks but *not* with their anticipatory eye gaze in an indirect AL task. This finding supports the idea that the development of an implicit ToM - in contrast to explicit ToM - is language independent.

In summary, Chapter 2 raised three issues that are important for the current thesis. First, different means of measurement reveal different onsets of FB understanding. Studies using children's anticipatory looking behavior as an indirect measure of FB understanding provide evidence for an implicit ToM already in 2-year-olds. In explicit FB tasks that require a verbal prediction children start to succeed around their fourth birthday. Second, findings on the impact of the linguistic nature of explicit FB tasks are not consistent in the ToM literature. Some studies suggest that an additional verbal input makes the task easier compared to a completely nonverbal task but there are equally studies that show a negative influence. Third, it is still under debate if children's early success in predictive looking tasks reflects the same ability as measured in explicit FB tasks. The role of language has received a lot of attention as in several theoretical accounts language acquisition is proposed to be relevant only for explicit but not for implicit FB reasoning.

The next chapter reviews the literature on the relationship between language acquisition, especially syntax, and the development of a ToM.

---

<sup>6</sup>In the training phase, children received verbal explanations including mental state language about the outcomes of FB stories.



# Chapter 3

## Language & ToM

‘‘They don’t know that we know that they know.’’

(Monica, taken from *Friends*, series 5, episode 14)

During the past 15 years, numerous studies<sup>1</sup> testing a variety of populations (typically developing children, specific language impaired children, deaf children, children with autism) have shown that children’s explicit FB understanding is associated with their linguistic abilities (for a review see Milligan et al., 2007, and Astington & Baird’s book *Why language matters for Theory of Mind*, 2005b). Researchers are holding different views about the nature of that developmental relationship. Some authors argue that the correlation found between language measures and FB task scores rather reflect the fact that most FB tasks are language demanding and any relation between both skills is thus a by-product (Chandler et al., 1989; Hahn, 2009), or that FB understanding rests on general cognitive operations that again require language to implement them (e.g., Fodor, 1992; Bloom and German, 2000).

Others, however, have claimed that language essentially contributes to ToM development and that the relation between both domains is fundamental (e.g., de Villiers, 2005a; Nelson, 2005). Experimental evidence speaks in favor of this latter approach, but even within this line of research there is an ongoing debate about the question which specific language property might promote ToM development. In their meta-analysis including 104 studies, Milligan et al. (2007) analyzed the effect of general language, semantics, receptive vocabulary, syntax and sentential complementation and found that all these aspects of language ability are related to FB understanding. However, receptive vocabulary was less strongly related to ToM performance compared to general language measures or syntactic abilities. The

---

<sup>1</sup>The search procedures in Milligan et al. (2007) revealed 324 studies investigating the relationship between both abilities until then.

authors yet acknowledge that relatively small effect sizes for vocabulary size are possibly due to the fact that receptive vocabulary tests tap into a language ability which is least overlapping with the other aspects of language.

In terms of its function, language has both a communicational, interindividual and a representational, intraindividual aspect (e.g., Astington and Baird, 2005a). Some researchers have proposed that children become aware of other's mental states through communication with others, they "enter the community of minds" through conversation (Nelson, 2005, p.32). This body of research stresses the role of the (parental) language input (Dunn and Brophy, 2005; Symons, 2004). Positive correlations were found between preschoolers' FB task performance and both their mothers' use of mental state terms (Ruffman et al., 2002; Slaughter et al., 2007) and their conversational experiences with siblings and friends (Brown et al., 1996). Regarding the communicational aspect of language, others have argued that the acquisition of the *semantics* of mental state verbs is crucial for ToM development (e.g., Olson, 1988; Peterson and Siegal, 2000). Thereby children are expected to acquire the concept of mental states when they encounter the semantics of lexical items (e.g. know, believe, remember) that refer to internal, nonobservable states.

This dissertation, however, focuses on the second approach which emphasizes the role of language as an internal system. Such an internal system assumedly allows children to form a representation of how a character in a story might think of a given situation. This line of research has claimed that it is the syntax, the structure of language, which assists FB understanding (e.g., Astington and Jenkins, 1999; de Villiers and de Villiers, 2000; Tager-Flusberg, 1997). In this context, the role of sentential complements has received a lot of attention. In a strong version of a *linguistic determinism* account that has been put forward by Jill de Villiers and her colleagues, ToM development depends on the acquisition of the syntax of sentential complementation (de Villiers, 2000; de Villiers and de Villiers, 2000; de Villiers and Pyers, 2002; de Villiers and de Villiers, 2003; de Villiers, 2005a, 2007). According to her hypothesis, this syntactic structure has unique properties that allow for representing false propositions and thus solving FB tasks. The next section will outline the main ideas of this *syntax-first* hypothesis and gives an overview about experiments that have tested this hypothesis empirically.

### 3.1 Linguistic Determinism Hypothesis: Syntax first

Following the Language of Thought hypothesis (Fodor, 1975), thinking takes place in a language (as Fodor labeled it: Mentalese) and the linguistic expression is always as complex as its natural reference. Thinking about FBs, thus, must involve complex

internal representations. In the philosophical literature, mental states are referred to as *propositional attitudes* since they express the mental posture of a person towards a proposition (Quine, 1956). Linguistically, propositional attitudes are denoted by complement sentences, that is a sentence that embeds a clause as its object like in (1):

- (1) Anne thinks that the cake is in the green box.

The content of the mental state is expressed in the complement clause that follows the optional complementizer *that* (*The cake is in the green box.*). This embedded proposition may be true or false<sup>2</sup>. The mental attitude the subject holds to the proposition is expressed by the verb of the matrix clause ('think'); the person who is ascribed to this mental state is the subject of the main clause (Anne). Crucially, the complete sentence remains true even the embedded proposition is false.

To solve a FB task, a child must be capable of separating reality from another person's false representation of reality, that is the child has to coordinate two different representations of a single situation. Language, specifically sentential complementation provides means for representing a proposition that is counterfactual but still true and action-guiding for another person. In other words, language allows us to represent the propositional content and beyond that, the attitude of a person towards it (Rozeboom, 1972). In their linguistic determinism approach, de Villiers and her colleagues claim that the mastery of the syntax of complement taking verbs with embedded sentential complements is a prerequisite for FB understanding:

“The language for discussing mental events provides the child with a formal means of embedding propositions, and thus provides a necessary ingredient for representing false beliefs.” (de Villiers and de Villiers, 2000, p.196)

In other words, thinking about a FB of another person requires a propositional complexity that can be expressed only by sentential complementation. According to de Villiers (1999; 2005a), it is in particular the mastery of the syntax of nonfactive communication (like *say, tell*) or mental state verbs (like *think, assume*) that enables children to reason about their own and others' mental states. De Villiers (1999) hypothesizes a feature in the child's grammar to be responsible for marking the complements of mental state and communicative verbs as not obligatorily true. Before children have acquired this feature, they perceive every sentence with a complement clause as factive in the sense of Kiparsky and Kiparsky (1970), that is they require the complement clause to be true itself. As a result, no FB reasoning is possible:

---

<sup>2</sup>This is the case in nonfactive complement-taking verbs like e.g. belief, think etc., see Chapter 4.

“An individual with less language or no language would not be able to formulate the appropriate representation of another person holding a false belief, and hence have no basis for reasoning about their actions. Such an individual would not only fail to give adequate explanations of why someone acted strangely (lacking the terms in which to express it to our satisfaction), s/he would also fail to show evidence of understanding why that person acted the way he did. In other words, no prediction would be possible.” (de Villiers and de Villiers, 2003, p.338)

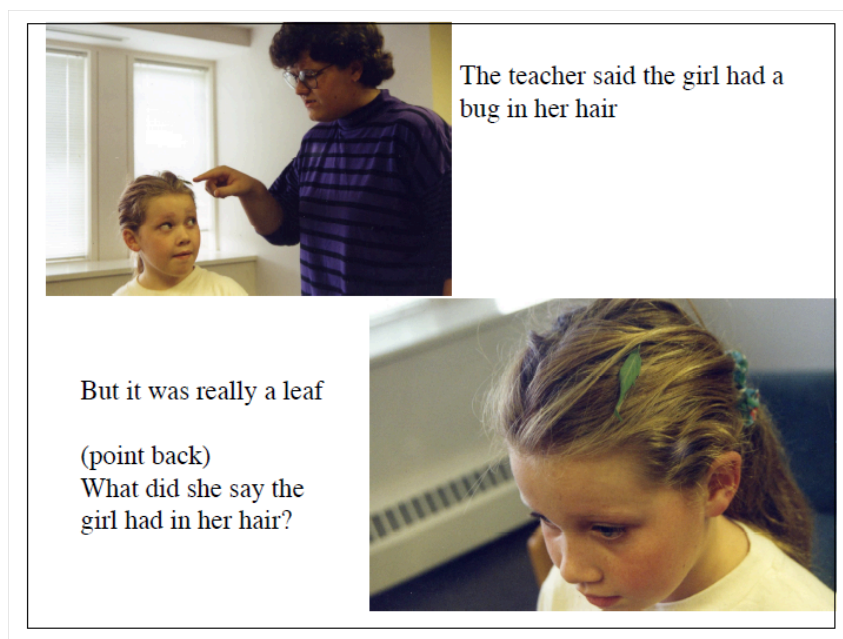
Other researchers and theorists have formulated similar notions about the relationship between language and ToM. Bickerton (1995) for example stated that language enables human beings to represent past, future and hypothetical events and hence to go beyond their own perceptual experience. Pinker and Bloom (1990) argue that the mechanism of complementation ”allows the expression of a rich set of propositional attitudes within a belief-desire folk psychology.“ (p.720). And more recently, Hauser et al. (2002) suggested that non-linguistic abilities like number cognition, navigation through space and ToM may rely on *recursion*, the component of the faculty of language that the authors identified as the only uniquely human one.

### 3.1.1 Empirical Evidence

**Typical development** The linguistic determinism hypothesis is supported by studies employing different kinds of experimental designs and testing different populations. First, cross-sectional studies have shown that the use of mental state verbs and their complements coincides in time with successful ToM task performance in typical developing preschoolers (de Villiers, 1995; de Villiers and Pyers, 1997).

Results of longitudinal studies that are qualified to determine the direction of this relation indicate that children’s early general syntactic abilities (Astington and Jenkins, 1999) and more specifically children’s comprehension of sentential complements predict later FB task performance, but not vice versa (de Villiers and Pyers, 2002). It is interesting to note that Astington and Jenkins (1999) consider general syntax and not sentential complements in particular to be the crucial device for solving ToM tasks in that it helps to keep track of and represent complex FB stories. However, they did not test children’s mastery on complement construction specifically. For testing this ability, de Villiers and Pyers (2002) devised their so-called *memory-for-complement task*. Each item of this task consists of two pictures (see an example in Figure 3.1) and a sentence like *The teacher said that the girl has a bug in her hair, but it was only a leaf*. The experimenter then points back to the first picture and asks the child *What did she say?* or *What did she say the girl has in her hair?*, respectively. In their longitudinal study, six out of twelve items included verbs of communication (*say, tell*), the other half contained mental state





**Figure 3.1:** Example of an item from the memory-for-complements task devised by de Villiers and Pyers (2002).

verbs (*think, belief*). Results revealed that only about 25% of 28 children aged 3;5 years on average (range 3;1-3;10 years) passed the task (i.e. they correctly answered *a bug*, or alternatively *that she has a bug in her hair*); three-quarter of the participants still answered incorrectly (*a leaf / that she has a leaf in her hair*). At a mean age of 3;8 years about 58% of the children passed this task and about 90% of the children passed at age 4;0 years<sup>3</sup>. The authors reported no differences between the two types of question (complex vs. simple question) which supports their idea that it is not an issue of processing limitation regarding complex questions but rather a matter of whether that hypothesized feature in the child's grammar is set or not.

Schulz and Ludwig (2008) reported similar results for German-speaking children: 12 out of 15 participants (80%) aged 4;2 years on average passed a translated version of the memory-for-complement task including eight sentences with the communication verb *say*. Crucially, none of the three failers displayed FB mastery.

<sup>3</sup>Criterion for passing was set at 10 or more correct answers out of 12.

The basic idea of the memory-for-complement task is that the child is only able to produce grammatical structures that s/he can represent and thus has acquired. However, this task has been discussed critically in the literature. Adler (2002) and Ruffman et al. (2003) claim that tests tapping the comprehension of complement constructions like the memory-for-complement task measure the same ability as measured in ToM tasks, namely to handle the fact that there can be two representations of one situation at the same time - a false one (i.e., the content of someone's mind or speech) and a true one (i.e., the true state of affairs). Hence, they conclude it is not surprising that de Villiers and colleagues find a correlation between these two scores. De Villiers and Pyers, however, claim that "This task does not require the child to 'read' the character's state of mind, but merely to represent it by holding the sentence in mind and then repeating the relevant piece back." (2002, p.1043). Yet, Flavell et al. (1990, Exp.3) found that 2;10 to 3;6-year-old children experience considerable difficulties in tasks in which they are required to reproduce counterfactual propositions. Only a third of them correctly repeated what a character mistakenly guessed about the color and size of a hidden object (e.g., *Hmm, I think you have a white cup over there.*; in reality the cup was blue). Notably, the character was not holding a FB about the object's property (as it is the case in de Villiers' task) but simply made a guess. Still, the children did not apply a simple parroting-strategy indicating that it is not a trivial task for 3-year-olds to repeat the false clause.

Two training studies demonstrated that teaching the syntax of sentential complements was sufficient to facilitate FB understanding (Hale and Tager-Flusberg, 2003; Lohmann and Tomasello, 2003) and are thus supporting the causality assumption. Hale and Tager-Flusberg (2003) trained 3-year-olds who failed at traditional FB task and the memory-for-complement task in one of three conditions: training on (1) FB reasoning, (2) sentential complements including verbs of communication or (3) relative clauses. Relative clauses were included as a control condition because of their structural similarity to complement clauses: both constructions allow for embedding propositions, but unlike sentential complements, relative clauses are embedded after nouns and are not obligatory. In addition, relative clauses have the same truth value like the main clause. After three to five days, all children were tested on FB tasks, complement clause and relative clause comprehension tasks. All children showed training effects within the tasks on which they were trained, and more importantly children in the complement-training group showed a transfer effect to FB reasoning. In contrast, training on relative clauses had no effect on the later ToM task performance and training on FB reasoning did not result in improvement on the complement comprehension task. This result contradicts results found in a study by Smith et al. (2003) who reported a positive correlation between both complementation *and* relative clause comprehension and FB task performance. Smith et al. (2003) argue that the mastery of two representations (i.e. the relative clause event that is embedded into a matrix clause event) is crucial for solving the

FB task, and the truth value of the embedded proposition does not play a role.

Lohmann and Tomasello (2003) aimed to find out which aspect of training is crucial to improve ToM task performance. They trained 3-year-old children who did not solve the FB task yet. In four sessions children were playing with deceptive objects (e.g. a pen that looks like a flower) and talked about it with an experimenter. Participants were randomly assigned to one of four linguistic conditions: (1) the experimenter commented on the objects using mental state or communicative verbs<sup>4</sup> in sentential complement constructions and highlighted the deceptive context (full training); (2) the experimenter used only sentential complements after mental state/ communication verbs but without mentioning the deceptive aspect (sentential complement only); (3) the experimenter highlighted the deceptive context but without using mental/communication verbs and sentential complement (discourse only training) or (4) the experimenter highlighted the deceptive aspect nonverbally with attention getter like *Look!* (no language training). Children who got the full training showed the highest improvements in an unexpected-content task; the no-language group did not improve at all. The other two groups increased their FB task performance slightly. This pattern of effects strongly suggests first of all that the use of language was necessary for improvement. Specifically, training on solely sentential complements as well as discourse without the use of mental language could independently enhance children's FB reasoning, but incorporating both factors appeared to be the most effective way to enhance FB understanding. In addition, there was no difference between communication and mental state verbs indicating that it is the syntactic construction and not the semantics of the used verbs that facilitates the FB reasoning. At first sight, the results provide evidence for the importance of language in ToM development, but as pointed out by Harris (2005), they also question its causal role. For Harris, the fact that improvement also took place in the discourse-only group demonstrates that sentential complementation is not a prerequisite for ToM development.

Further insights about the connection between ToM and language stem from studies with adult subjects. It seems to be the case that language is not only important for the development of a ToM but also for the computations necessary during ToM task performance. Using a dual-task paradigm, Newton and Villiers (2007) examined adults' performance on a nonverbal FB task while they were either simultaneously performing a verbal interference task (verbal shadowing) or a nonverbal interference task (rhythmus shadowing). FB reasoning was disrupted in the former, but not in the latter condition indicating that language is needed for the access to (at least explicit) FB understanding. However, adults with severe aphasia who display impairments even in comprehending non-embedded sentences have been reported to succeed on FB tasks (Siegal et al., 2001; Varley and Siegal, 2000; Varley et al., 2001). This

---

<sup>4</sup>Half of the group was trained on mental state verbs, the other half on communication verbs.

suggests that thinking about mental states can take place without the assistance of language and ToM is - at least after it is in place - independent from linguistic resources.

**Atypical development** There are three atypical developing populations that provide insightful information regarding the relation between language and ToM: individuals with autism spectrum disorders (ASD), children with specific language impairment (SLI) and late-signing deaf children/adults.

Autism is a pervasive developmental disorder which is characterized by impaired communication, impaired social interaction and repetitive behavior (Dilling et al., 2009). According to the ToM-hypothesis, impairments of ToM are assumed to explain deficits associated with ASD, especially impairments in communication and social interaction (Baron-Cohen et al., 1985; Frith, 1989). And although children with ASD do perform consistently worse on FB task than age-matched controls (see Baron-Cohen, 2000, for a review) some children pass (e.g., 4 out of 20 participants passed in Baron-Cohen et al., 1985).

Happé (1995) reported a close relation between language ability and FB understanding in individuals with ASD. In a longitudinal study, Tager-Flusberg and Joseph (2005) showed that general syntax and complement clause comprehension uniquely predicted concurrent and later ToM task performance in children with ASD. They used an enhanced version of the memory-for-complement task: besides false complements following communication and mental verbs they included also true complements. This was done in order to examine both the semantic and the syntactic property of sentential complements. Whereas comprehending true complements requires only syntactical knowledge, false complement comprehension requires syntactic *and* semantic knowledge because of the conflict between the embedded proposition and true state of affairs. Interestingly, only the mastery of false complements with communication verbs contributed uniquely to the variance in later FB scores; cognition verbs and communication verbs with true complements did not account for additional variance. Tager-Flusberg and Joseph (2005) suggest that verbs of communication may provide a bootstrap for FB understanding in children with autism: since verbs of communication involve the same syntactic structure like mental verbs and they offer a way to express false propositions, children with ASD might learn to build an analogy to mental phenomena and thus reason logically in ToM tasks without having necessarily a conceptual FB understanding. This kind of compensatory learning is supported by studies measuring spontaneous, implicit responses in autistic adults and children: whereas they did not display any implicit FB understanding in their eye gaze they performed better in a FB task in which they were requested to make explicit predictions (children: Ruffman et al., 2001; Hahn, 2009; Senju et al., 2010; adults: Senju et al., 2009). Children's verbal FB

task performance was correlated with their general language ability (Ruffman et al., 2001b) and complement clause comprehension (Hahn, 2009). According to Ruffman et al. (2001), individuals with ASD are fundamentally impaired in their social understanding as indexed by their lacking spontaneous correct looking behavior but they are - as language ability increases - able to solve FB task explicitly.

Children with SLI form the second atypical developing group that provide important information about the relation between language and FB understanding. Individuals with SLI are delayed in their language development but without having a cognitive or neurological deficit (Leonard, 2003). Studies examining the FB understanding in SLI-children found that these children are also considerably delayed in their ToM development (de Villiers et al., 2003; Gillott et al., 2004; Farrant et al., 2006; Farrar et al., 2009; Hahn, 2009; Miller, 2001). This supports the hypothesis that language plays a causal role for ToM development. To examine further whether this delay is due to the linguistic demands of traditional FB tasks, Hahn (2009) presented both a verbal and a nonverbal FB task to 5;0 to 7;7-year-old SLI-children. Participants did not perform above chance in both tasks and no difference between the two versions was found. Miller (2001) presented four different FB conditions that differed in their linguistic complexity. In contrast to Hahn (2009), SLI-children performed similarly to normally developing controls but only in reduced linguistically demanding FB tasks. This indicates that ToM ability in SLI-children is masked by the linguistic nature of traditional FB tasks. Modifying her own former results, Miller (2004) added a less-verbal condition (i.e. showing silent FB movies) to three previous used conditions and found that age-matched controls and SLI-children did not differ in their performances in any of the four conditions. This study also included children's comprehension of sentential complementation as a predictor of FB task performance. Interestingly, SLI-children performed significantly worse than age-matched controls on a modified memory-for-complement task, but still they were able to pass two out of four FB conditions. On the one hand, this pattern of results speaks against the syntax-first hypothesis. On the other hand, significant correlations were found between complement mastery and FB scores in both SLI and normally developing children.

De Villiers et al. (2003) also systematically investigated the contribution of sentential complement comprehension to FB understanding in SLI-children and found besides a general ToM delay that complement comprehension significantly predicted FB task performance after controlling for age-effects. This did not hold for other language measures like morphosyntactical knowledge or general speech processing skills. Contrasting results were obtained by Swoboda (2006) and Farrar et al. (2009). Swoboda (2006) tested 18 SLI-children and 42 typically developing controls and found no relation between their language proficiency status and composite ToM scores. Although SLI-children performed considerably worse on complement comprehension

tests compared to controls, some of them were able to solve FB tasks. However, significant correlations were found between children's FB understanding and their vocabulary size, sentence memory and phonological knowledge<sup>5</sup>. Similarly, Farrar et al. (2009) found that general language measures (productive vocabulary, expressive syntactical ability) significantly predicted concurrent FB task performance but the comprehension of sentential complementation did not make a unique contribution to ToM scores. In summary then, although results are rather mixed, there is strong empirical evidence that SLI-children are delayed in their ToM development because of their delayed language acquisition.

Like SLI-children, deaf children with hearing parents (so called late-signers) are, compared to children who are exposed to sign language from birth on, delayed in their language acquisition but not impaired in their non-verbal cognitive development. Studies that have examined FB understanding in late-signing children have found a delayed ToM development relative to hearing controls or native signers (de Villiers, 2005b; Gale et al., 1996; Peterson and Siegal, 1999, 2000; Peterson et al., 2005; Schick et al., 2007; Woolfe et al., 2002). This is not only a simple reflection of their language comprehension limitations in verbal FB tasks since they are also delayed in less-verbal or nonverbal ToM tasks (Gale et al., 1996; Figueras-Costa and Harris, 2001; Schick et al., 2007). Schick et al. (2007) demonstrated that one of the highest predictors of FB reasoning in language-delayed deaf children is their understanding of complements. De Villiers (2005b) reported that besides complement clause comprehension also vocabulary size was an important predictor for deaf children.

A special group within the deaf population are the signers of the Nicaraguan Sign Language (NSL). They provide compelling evidence that human language abilities are a prerequisite of FB understanding. Nicaraguan deaf individuals started to develop a completely new sign language at the end of the 1970s when they were allowed to join a special school for deaf children; before they had no or just little contact with each other (e.g., Senghas, 1995, 2003). The first cohort of children started with simple gestures and had (and still have) difficulties in reasoning about FB, whereas children of the following cohorts possess more complex linguistic skills and are able to pass FB tasks (Morgan and Kegl, 2006; Pyers and Senghas, 2009). This performance gap is assumed to be caused by lacking early conversational experience. However, Pyers and Senghas (2009) showed recently in a longitudinal study that this gap between late and early NSL signers decreases with time. Participants of the first cohort learned to use mental state verbs (measured with an elicitation task) and also improved their ToM task performance (assessed in a nonverbal test). The authors conclude "that, with the increasing contact, first-cohort signers were exposed to a

---

<sup>5</sup>Note that Swoboda (2009) did not partialize out age in these analyses, that is it could simply reflect an age effect.

form of NSL that was richer than their own and that included the new mental-state words produced by their younger peers.” (Pyers and Senghas, 2009, p.810). It is important to note, however, that results of studies with NSL signers are compatible with both the metaawareness-through-conversation account and syntax-first account.

In summary, several studies testing typically and atypically developing children provided evidence that the comprehension of sentential complements plays a causal role for the development of FB understanding. It is still an unsolved question of whether the role of language is only in the ToM development or whether it might also be in the on-line computation during FB reasoning. Yet, the proposal that ToM development can not proceed without a specific syntactic understanding is discussed controversially. The following section gives an overview about the criticism of the linguistic determinism hypothesis and about alternative accounts that have been put forth.

### 3.1.2 Criticism and alternative accounts

As mentioned earlier, one main criticism of the syntax-first hypothesis concerns the method of how children’s comprehension of sentential complements has been tested so far (Adler, 2002; Ruffman et al., 2003; Ruffman, 2004). The memory-for-complement task requires the handling of misrepresentation to some extent. According to Adler (2002) and Ruffman et al. (2003), this task is in fact a FB task with reduced task demands since it does not (as standard FB tasks do) require the child to follow a complex narrative and to predict an upcoming action. All studies providing evidence for the syntax-first hypothesis have used this task or a modification of it. Lohmann and Tomasello (2003) even used Swettenham’s ToM test (1996) to measure participants’ sentential complement comprehension<sup>6</sup>. This demonstrates how closely linguistic and cognitive abilities are related. Thus, language tests used to examine the linguistic determinism hypothesis need to suitably distinguish possible interactions. Hence, the most crucial test for the comprehension of tensed complements without concurrently tapping the understanding of (mental) misrepresentations would thus involve true complements and non-mental complement taking verbs.

Studies which found that not complementation mastery but rather general language skills are predictive of FB understanding challenge the linguistic determinism hypothesis from a more theoretical point of view. In a longitudinal study, Ruffman et al. (2003, Exp.1) found that semantic but not syntactic abilities uniquely predicted later FB understanding. It is important to note that only word order mastery was taken as a syntax measure in this analysis. In Experiment 2 (cross-sectional), they

---

<sup>6</sup>Children were required to predict whether a story’s character will take his raincoat after they heard that e.g. “The boy thinks that it is sunny outside, but it is raining.”

included relative clause comprehension and did find correlations between embedded clause comprehension *and*, again, semantic abilities. Interestingly, Slade and Ruffman (2005) reported a bidirectional relationship, that is early FB task scores also predicted later language test performance. Thus, the authors concluded, both semantic and syntactic development contribute to FB understanding and there is no causal relationship between language and ToM. Similarly, Tardif et al. (2007) and Cheung et al. (2004) found children's general language competence to be best predictive of FB understanding while complement comprehension did not contribute uniquely to the variation of childrens' ToM task scores.

A third line of criticism arises from studies that examined the acquisition of sentential complementation in other languages and show cross-linguistic differences. As for German, Perner et al. (2003) investigated the verb of desire (*want*) (see also Perner et al., 2005). In German, a desire can be expressed with an infinite complement when it concerns the subject's action, analogous to English like in (2a). However, a finite complement phrase has to be used for expressing more complex desires as in (2b). This is ungrammatical in English.

- (2) a. Anne möchte Süßigkeiten kaufen.  
       Anne wants to buy some sweets.  
       b. Anne möchte, dass ihre Mutter Süßigkeiten kauft.  
       \*Anne wants that her mother buys some sweets.

Although verbs of communication and cognition appear in the same syntactic structure as verbs of desire, German 3- to 4-year-old children performed considerably better when they were asked to remember complements after desire statements compared to complements following verbs of belief or communication. However, German children also passed FB tasks around the age of 4 and not earlier<sup>7</sup>. This strongly suggests that it is not solely the syntactic knowledge of tensed complements which boosts ToM development. Tardif and Wellman (2000) reported a similar acquisitional pattern for Cantonese- and Mandarin-speaking children's production of *want* and *think*: *want* emerges well in advance of *think* although both verbs take the same syntactic form<sup>8</sup>. The authors conclude that these results contradict the assumption that tensed complements alone provide a mental structure for thinking about FB. Instead they provide evidence for a conceptual change that takes place along the desire-belief continuum and which has nothing to do with the linguistic expression of the concepts per se. de Villiers (2005a) commented on

<sup>7</sup>Unfortunately, Perner et al. (2003) did not report any correlations between language measures and FB task performance.

<sup>8</sup>Note that compared to English both Chinese languages have a simpler grammatical constructions after mental state verbs.



this critique by pointing out the subtle differences between the syntactic forms of *want*- and *think* statements: desired events always refer to hypothetical events, that is the complement clauses are *irrealis*; there are no false desires. The embedded proposition following *want* is discrepant with reality but there is no truth value on the complement that can be evaluated. De Villiers (2005a) thus restricts her linguistic determinism hypothesis to *realis* tensed complements.

Cantonese is also interesting for the following reason: here the belief status can be coded either as neutral (*nam5* means ‘to think’) or as false (*ji5wei4* means ‘to think falsely’). Tardif et al. (2004) investigated whether children acquiring Cantonese display a different developmental trajectory in their FB understanding because of this specific marking in their language. They found that mental verb semantics did have an effect: when the belief was explicitly marked as false (*ji5wei4*) children performed better in a FB task than when the neutral word for think was used. The general developmental pattern, however, was found not to be different from English-speaking children in that sense that Cantonese-speaking children do not display a considerably earlier FB understanding. Cheung et al. (2009) investigated also the lexical impact of different mental verbs in Cantonese and found that while verb semantics (especially the nonfactive *ji5wei4*) predicted ToM task performance, complement syntax alone had no effect.

Perner et al. (2005) shares the central assumption with de Villiers’ theory, namely that during development children learn to represent different perspectives which enables them to solve FB tasks. Yet, whereas de Villiers assumes that solely the linguistic system provides the device by setting a *Point of View* feature in the grammar, for Perner language as a communication system supplies children with the information they need to construct a ToM; a special role of syntactic competence is not considered. In other words, the ability to manage different perspectives is rather reflected in children’s FB task performance *and* their language abilities. Studies testing children’s metalinguistic awareness and FB understanding support this approach: a positive relationship between childrens ToM task performance and their understanding of homonyms (Doherty, 2000) and synonyms was found (Doherty and Perner, 1998; Perner et al., 2002). In this context, it is interesting to note that bilingual children who are experienced in having different names for the same referent outperform monolinguals on ToM tasks (Kovács, 2009; Goetz, 2003). This cannot be explained by bilinguals’ better language skills as Kovács (2007) reported that mono- and bilinguals do not differ in their acquisition of complement clause comprehension.

Another possible route of ToM development is suggested by the following results of a training study (Lohmann and Tomasello, 2003; Lohmann et al., 2005): the conversation about deceptive objects without using sentential complementation and mental state terms already enhanced children’s FB task performance. Thus, it is possible that any discourse about the discrepancy between an object’s appearance and its real function helps the child to build up a ToM. De Villiers (2005a), however,

noted that the improvement in this training condition was quite small and that it is not clear whether there were children who did not understand FB without understanding complement sentences, too. This would be the litmus case against the linguistic determinism hypothesis. Yet, this was reported in a study by Hahn (2009): there were children who solved a verbal FB task, but not the complement task. Moreover, tested with a nonverbal FB task, no correlations between language measures (including complement clause comprehension) and ToM were found.

In summary, the major criticisms of the linguistic determinism hypothesis concern (1) the memory-for-complement task which has been claimed to tap into children's FB understanding and (2) the specific impact of complex syntax on the ToM development as other aspects of language have been found to promote FB understanding as well. Furthermore, there is evidence that children sometimes pass the FB task without demonstrating an understanding of sentential complementation. The early manifestation of an implicit FB understanding in predicative looking tasks challenges the linguistic determinism hypothesis, too. The next chapter describes one experiment that has been conducted to investigate whether complement mastery also correlates with implicit FB understanding. This would be an indication that a specific syntactic understanding provides the critical means for representing FBs in *any* kind of FB task.

## 3.2 Implicit FB understanding & Language

The question arises of how it is possible that a child might exhibit adequate looking behavior and thus indicate FB understanding while not understanding sentential complements. Does implicit ToM operate independent of language? And if so, what are the specific mechanisms underlying these two levels of FB task performance? De Villiers and de Villiers (2003) assume that language might play a role only in direct ToM tasks, that is whenever it is required to reflect on propositional attitudes explicitly. A task measuring visual orientation is on the contrary likely to tap knowledge which is probably "based on empathy, or simulation of what I myself would do if faced with such a set of circumstances" (de Villiers and de Villiers, 2003, p.336). Ruffman (2004) suggested that implicit understanding of belief might be based on the (unconscious) detection of behavioral regularities and thus may be independent from language whereas in tasks in which a conscious decision is required, language (for Ruffman, it is both semantics and syntax) provides the means for reflecting on FBs explicitly.

To test this assumption, Low (2010) investigated the relationship between complement mastery and implicit FB reasoning in 3- and 4-year-olds. In experiment 1, he measured children's anticipatory eye gaze while they were watching a FB and a

TB scenario which were enacted live with puppets in a card house (similar to Ruffman et al., 2001a) and conducted a memory-for-complement task (analogous to Hale and Tager-Flusberg, 2003) with 4 items including false propositions after a verb of communication (*say*). Besides, age, vocabulary size and nonverbal intelligence were measured as control variables. Explicit ToM was tapped by verbal answers to direct FB questions in three different FB tasks. Results show that explicit, but not implicit ToM task performance<sup>9</sup> was related to both language measures, i.e. complement mastery and vocabulary size. In addition, both implicit ToM task and complement task scores uniquely predicted explicit ToM task performances after controlling for the other variables. In experiment 2, he included two low-verbal, but direct ToM tasks (thought-bubbles task derived from Woolfe et al., 2002, and object-hiding task from Call and Tomasello, 1999). The results indicate that any type of explicit FB understanding in terms of making a conscious decision seems to be linked to language irrespective of the linguistic nature of the FB task. In order to control for timing and stress, children were presented with a movie instead of live experimenters in a third experiment. This modification had no impact on participants' eye gaze.

Altogether, Low's results suggest that implicit FB understanding is language independent and - together with complex syntactic knowledge - is a prerequisite for explicit FB reasoning, at least in typical development. Low (2010), however, acknowledged that further research is necessary to define the specific role of complex syntax because the complementation task involved only false embedded propositions. That is why "It will be crucial to develop other viable analogues of the complement mastery task to delineate the extent to which it is the implications of certain syntactic constructions that partly lead children to explicitly understand differences in minds." Low (2010, p.613).

---

<sup>9</sup>The looking time difference between the object's assumed and real location in the FB movie was used here as metric for implicit ToM.



## Chapter 4

# The Acquisition of Sentential Complementation

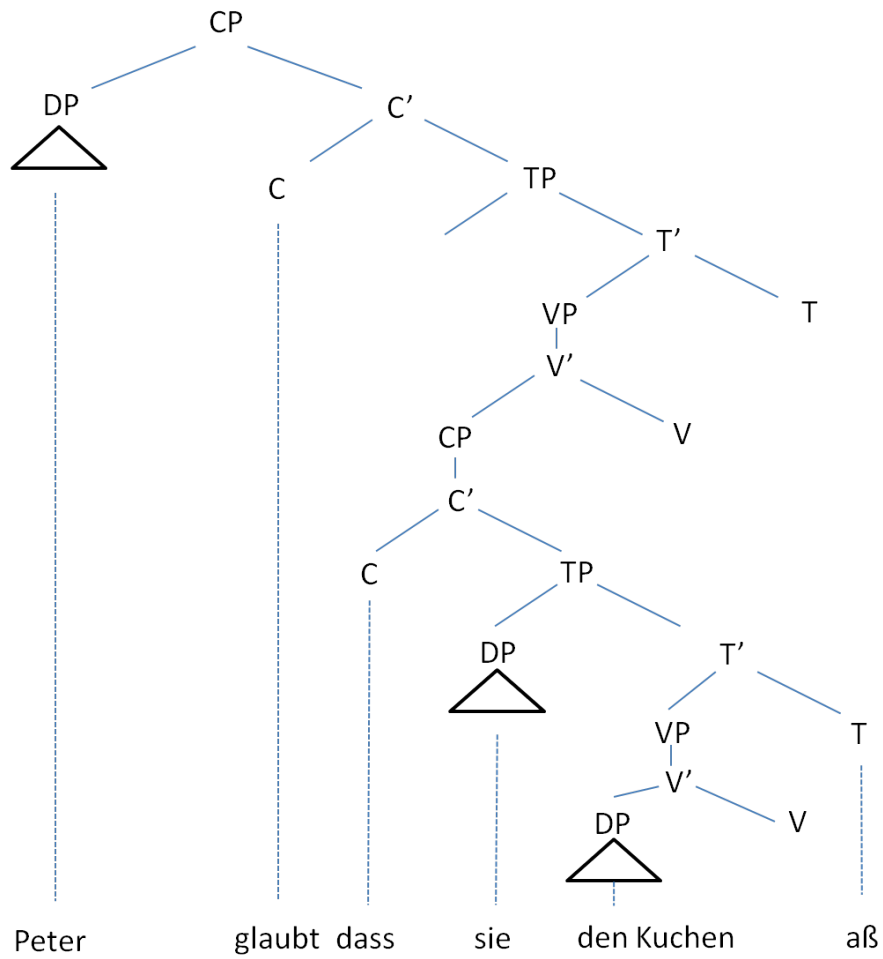
Language-acquiring children have to come to comprehend and produce utterances of different syntactic complexity, starting with single words and phrases, simple sentences up to complex sentences containing subordinate clauses (e.g., Clahsen, 1988). Sentential complementation marks an important type of syntactic complexity because the subordinate clause functions as an argument of a complement-taking verb (henceforth CTV) (e.g., Noonan, 1985). Since CTVs include a certain set of verbs, for example mental state verbs (*think, know*), both lexical and syntactic development are closely linked in this area providing an insightful way to investigate their interactions (e.g., Gleitman, 1990; Naigles, 2000).

This chapter reviews research on how children acquire complements under CTVs. The first section describes the properties of complement sentences and CTVs. In the following, empirical studies are presented. There are two sources of evidence on the acquisition of sentential complementation: investigations of the production of complement clauses in children's spontaneous speech corpora and experiments testing children's comprehension of sentential complementation. Section 2 summarizes acquisition data from production studies and section 3 discusses comprehension data.

### 4.1 Properties of Complement Sentences and CTVs

First, the terminology of the phenomenon of interest will be established. The construction type I refer to as a *complement sentence* is exemplified in (1). The italicized constituent in (1) is called the *complement clause*, in this case it is the sentential object of the CTV (*to think*), and the remaining part is called the *matrix clause* (see Noonan, 1985; Boye and Harder, 2007 for similar definitions). Figure 4.1 displays the syntactic configuration of the German complement sentence (1).

- (1) GERMAN Peter glaubt, *dass sie den Kuchen aß.*  
 ENGLISH Peter thinks *that she ate the cake.*



**Figure 4.1:** Simplified syntactic tree of the complement sentence: *Peter glaubt, dass sie den Kuchen aß.*

Complement clauses are thus defined as “special instance of complex sentences in which one proposition serves as an argument within another proposition” (Bloom et al., 1989, pp.101-102). While a complement clause is required by a CTV other subordinate clauses (e.g., relative clauses or adjunct clauses) are not obligatory. There are different types of complement clauses: nonfinite complement clauses (e.g., “Peter saw the child falling.“, see Quirk et al., 1985; Diessel and Tomasello, 2001) in which the verb is nonfinite and a complementizer is missing, and finite complement clauses (like in (1)) which have a finite verb and can be introduced by a complementizer. Besides, there are three types of complement clauses depending on the complementizer that is used: sentential complements that can be introduced by *that* like in (1), if-complements introduced by *if* or *whether* and wh-complements marked by a *wh*-pronoun or *wh*-adverb (Diessel and Tomasello, 2001). However, in the present work only finite sentential complements are considered because of its assumed causal role for the development of a ToM.

The use of the English complementizer *that* and the German equivalent *dass* is facultative. In contrast to English, German complement clauses that are introduced by *dass* are verb-final clauses (see Figure 4.1), whereas the embedded clause has verb-second word order when no complementizer is used (e.g., Rothweiler, 1993). German-learning infants have been shown to be sensitive to word-order violations in complement clauses: Using the Head-Turn-Preference Paradigm, 20-month-olds displayed different looking times to grammatical complement sentences (e.g., *Bert sagt, dass Lisa Oma hilft.*) compared to ungrammatical complement sentences (*Bert sagt, dass Lisa hilft Oma.*) (Weissenborn et al., 1998). Furthermore, Rothweiler (1993) and Brandt et al. (2010) reported that word-order errors in complement clauses are quite rare in German-speaking children’s spontaneous speech.

It has been proposed that the use of the complementizer is related to various discourse features and its presence may mark full-fledged complement clause constructions (for German children: Brandt et al., 2010; for English children: Diessel and Tomasello, 2001; for English adults: Thompson and Mulac, 1991). Concerning comprehension, de Villiers found no difference between sentences without complementizer and sentences containing *that* complementizer in English children’s performance in the *memory-for-complement task* (see 3.1.1 for detailed description of the task) (personal communication between Tager-Flusberg & Joseph and Jill de Villiers in 2002, reported in Tager-Flusberg and Joseph, 2005). This suggests that the additional marking of the subordinate status by a complementizer does not help English preschoolers to interpret the complex structure.

Noonan (1985) divides CTVs into different semantic classes:

- utterance verbs: *say, tell, report, promise, ask* etc.,
- propositional attitude verbs: *think, believe, suppose, assume, doubt* etc.
- pretence verbs: *pretend, make belief, imagine* etc.
- commentative verbs/ factives: *regret, be sorry, be sad, be odd, be significant, be important* etc.
- verbs of knowledge and acquisition of knowledge: *know, discover, realize, find out, forget* etc.
- verbs of fearing: *be afraid, fear, worry, be anxious* etc.
- desiderative verbs: *want, wish, desire, hope* etc.
- immediate perception verbs: *see, hear, watch, feel* etc.
- others (mostly used with the infinitive complement type)

Kiparsky and Kiparsky (1970) distinguish between *factive* and *nonfactive* CTVs. Factive verbs require the embedded clause to be true like in *I regret that Sally lost her money.*, i.e. in case Sally did not lose her money the entire sentence is infelicitous. Verbs of knowledge and commentative verbs belong to this class. In contrast, sentences with nonfactive verbs can be felicitous although the embedded clause is not true (e.g., in (1) in case she did not eat the cake). Propositional attitude predicates belong to this class since they express a positive (e.g. *believe*) or negative (e.g. *doubt*) attitude about the truth of the proposition, but in fact its truth value remains undetermined. A third class called the *counterfactives* comprising pretence and desiderative verbs (wish, pretend) imply that the embedded proposition is false at the time of the utterance (see also Harris, 1975; Schulz, 2003). Verbs of communication (i.e., Noonan's class of utterance verbs) are neither factive nor nonfactive as they allow, like propositional attitude verbs, false embedded propositions with full tensed sentences, for instance when expressing that someone lied or told something wrong.

According to de Villiers' linguistic determinism hypothesis (see section 3.1) it is the acquisition of nonfactive verbs with their potential to convey false propositions, which enables the child to represent mental states of others and thus develop a ToM. It is proposed that nonfactives have a syntactic feature in the complementizer position indicating that the following complement *can* be false (de Villiers, 1999, 2005a). This is also called the *Point of View* (PoV) feature which marks the truth of the complement clause as being relative to the subject of the matrix clause and not relative to the speaker. But how do children acquire this PoV feature? de Villiers



proposes a bootstrapping mechanism: Children interpret per default complement clauses as being true (de Villiers, 1999, 2005a). As language development proceeds, they assumedly learn first that verbs of communication can take false propositions and, in a second step, use the analogy of the verb classes by extending this knowledge to nonfactive mental state verbs. The acquisition of communication verbs is assumed to precede the acquisition of (nonfactive) mental state verbs because the propositions embedded under verbs of communication can be evaluated with respect to their truth as acts of speaking are overt whereas acts of thinking are not. For de Villiers (1999, 2005a), the understanding of nonfactivity constitutes a full-fledged comprehension of sentential complementation:

”[...] full mastery of mental verbs and their complements entails a step-wise development, with its final manifestation being a full syntactic structure with complements marked by a PoV.“ (de Villiers, 2005a, p.213)

A similar approach to explain the acquisition of mental state verbs was proposed by Papafragou et al. (2007). Assuming that word learning relies at first on mapping between what is seen in the extralinguistic environment and what is said (also called *word-world-mapping*), mental verb learning is a challenging task because the concepts they express are referentially not transparent. Gleitman and others propose a *syntactic bootstrapping* mechanism to explain how (mental) verbs are acquired (Fisher et al., 1991; Gleitman, 1990). The basic idea is that the unique syntax that is required by mental state verbs - namely sentential complementation - helps the child to narrow the potential meaning of a novel, mental lexical item<sup>1</sup>. In other words, the structure provides information about the semantics. For example, in a sentence such as (2) the listener must infer some kind of mental or communication activity of Matt about his grandmother.

(2) Matt GORPS that his grandmother is under the covers.

Papafragou et al. (2007) presented adults and 3- to 5-year-olds videotaped stories (Exp.3). At the end of each movie, one of the story protagonists summarized what happened in the story using a nonsense word in the verb position as in example (2). Participants were asked then to guess the meaning of a nonsense verb that was mentioned. Results indicate that both linguistic cues (clausal that-complements) and situational cues (a FB context) increased the likelihood to interpret novel verbs as mental state verbs.

Another important distinction for de Villiers' theory is the classification into *realis* and *irrealis* CTVs which codes the observability of the embedded proposition: *realis* indicates that something is actually the case (e.g. *know, say*); *irrealis* verb

---

<sup>1</sup>Note that this holds also for other CTVs like verbs of communication or perception etc.

forms are used when speaking about a future or hypothetical event (e.g. *wish*, *want*). In contrast to English, expressing a complex desire in German (e.g., to want someone else to do something) requires a complement sentence. Perner et al. (2003) tested German-learning preschoolers on their ability to extract complement clauses embedded under *want* and under *think* after they heard complex sentences. 3- to 4-year-olds mastered *want*-sentences considerably earlier than *think*-sentences although both share the same syntactic form: about 70% of the children aged between 2;5 and 3;7 passed the *want*-condition, but only about 10% of this age group passed the *think*-condition. Yet, the development of a ToM is not accelerated in German-acquiring preschoolers. Based on this finding, the authors argue against the linguistic determinism hypothesis and propose that the emergence of a ToM follows a certain developmental pattern irrespective of language development (see also Perner et al., 2005). de Villiers, however, emphasizes that complements after verbs of desire are always *irrealis*, that is the embedded proposition cannot be evaluated against reality. In other words, there is no false desire. Therefore, she restricts her syntax-first proposal to *realis*, nonfactive CTVs taking finite complement clauses (de Villiers, 2005a).

## 4.2 Production Studies

Corpus-based studies have shown that English- and German-learning children start to produce complement sentences from the age of 2 years (English: Bloom et al., 1989; Diessel and Tomasello, 2001; Diessel, 2004; Limber, 1973; German: Brandt et al., 2010; Rothweiler, 1993), and mental state terms emerge in children's speech from the second half of the third year (Limber, 1973; Shatz et al., 1983). The use of CTVs is dependent on their function: Children first refer to perception, emotion and desire, then to cognition, knowledge and belief (Bartsch and Wellman, 1995; Bretherton and Beeghly, 1982).

However, several authors pointed out that there are different functions of mental state verbs when used with sentential complements: a formulaic/ parenthetical use as in (3a<sup>2</sup>) where the main clauses can be "interpreted as clausal operators that modify the content of the rest of the sentence"<sup>3</sup> (Kidd et al., 2005, p.51) and an assertive use (3b) where the complement clause is unequivocally embedded and the whole sentence expresses two propositions (Diessel and Tomasello, 2001; Nixon, 2005; Kidd et al., 2005; Thompson, 2002).

---

<sup>2</sup>This example is taken from Kidd et al., 2005.

<sup>3</sup>The verb of the matrix clause is often in first person singular present active (Diessel and Tomasello, 2001).

- (3) a I think I want grape juice.  
b He thinks I want grape juice.

Limber (1973) already noticed a parenthetical use of ‘I think’ in children’s spontaneous speech and Shatz et al. (1983) also stated that “earliest uses of mental verbs are for conversational functions rather than for mental reference.” (p.301). 4-year-olds have been found to use mental state verbs both when referring to cognitive states and equally often as pragmatic markers (e.g., *It is a dog, I think?*) whereas sentential complementation was more likely to occur in mental referencing contexts (Nixon, 2005). Hooper and Thompson (1973) pointed out that even in adult English, verbs that can take complement clauses also have a parenthetical use.

In a longitudinal study, Diessel and Tomasello (2001) and Diessel (2004) analyzed the acquisition of sentential complementation in the spontaneous speech of seven English-speaking children from age 1;2 to age 5;2. They demonstrated that children’s early use of finite complement clauses is solely formulaic, that is the main clause serves rather as an attention getter, epistemic marker or marker of illocutionary force instead of expressing a discrete proposition. In their study, children started to produce bi-clausal complement sentences and therewith unequivocally embedded clauses from the age of 4 years with most of the verbs.

Diessel’s interesting results have been influential, but also have some limitations because the corpus data may not be informative about the onset of clausal complementation in child language. The database of CTVs that Diessel and Tomasello (2001) have used contained 1811 potential complement clauses. A study of spontaneous speech in adults has shown that in a corpus of 425 potential complement clauses in spoken English only one was unequivocally embedding<sup>4</sup> (Thompson, 2002; but see Newmeyer, 2010, for different results). Hence, the fact that truly embedded complement clauses are relatively rare even in adult spoken language suggests that analyzing the production of complex sentences might not reveal when children exactly start to understand complement sentences containing truly embedded clauses. Thus, the current work employed experimental methods to examine the onset of sentential complementation in German-learning children. The next section will provide an overview of studies testing children’s comprehension of complement sentences.

### 4.3 Comprehension Studies

Numerous studies investigating the understanding of CTVs and their complements examined when children start to understand mental state verbs including their associated (non)factivity status, that is whether the embedded complement clause is presupposed to be true (like for *know*, factive) or whether the truth value is

---

<sup>4</sup>Boye and Harder (2007) claim that the count really should be 3 out of 425.

undetermined (like for *think*, nonfactive).

The difference between *think* and *know* seems to be understood by the age of 4 (Johnson and Maratsos, 1977). 4-, but not 3-year-olds correctly confirmed that a misinformed seeker rather thinks than knows where a hidden object actually is, indicating an understanding of the different meanings of these mental state verbs. Hopmann and Maratsos (1978) investigated 3- to 7-year-olds' comprehension of (non)factive CTVs using a forced-choice task: Participants were required to select one of two possible actors to act out an utterance in which one character was mentioned as the subject of the complement clause. Crucially, if a nonfactive CTV is negated in the main clause (e.g., *It isn't possible that the girl bumps into the duck.*) the second, unmentioned agent has to be chosen (e.g., a boy). In contrast, the complement clause remains true if a factive CTV is negated (e.g., *He does not know that the girl bumps into the duck.*). Preschoolers were reported to choose the agent that was associated with the subject of the complement clause in both verb conditions. Thus, they assumedly process the complex sentence only partially, that is they tended to focus on complement clauses regardless of the negation condition (this was labeled as *overextended affirmation tendency*). This bias however, might reflect a task-specific strategy since children had to act out the complement clause.

Using a less demanding task, Schulz (2002) presented English-speaking children with complement sentences containing (1) factive (e.g. *find out*), (2) nonfactive (e.g. *think*) and (3) counterfactive CTVs (e.g. *forget*) and asked simple yes/no questions about the truth of the embedded proposition (example for a nonfactive item: *The boy thought that there was an ant in the bowl. Was there an ant in the bowl?*). Results showed that 4- to 6-year-olds correctly assigned a factive or nonfactive interpretation to the complement clause depending on the matrix verb (expected answers: *yes* for factives, *no* for counterfatives and *maybe/ don't know* for nonfatives). Schulz and Ludwig (2008) tested German-acquiring 3- to 4-year-olds (age range: 3;05-4;10) with the same task and found that children who showed FB understanding performed better in the nonfactive CTV condition than ToM task failers. Interestingly, even ToM task passers answered in 50% of the cases still incorrectly. Hence, Schulz and colleagues argue that FB understanding might be a prerequisite for the ability to interpret different presupposed truth values of complement clauses and thus to comprehend the (non)factivity status of a CTV (Schulz, 2003; Schulz and Meissner, 2008).

Abbeduto and Rosenberg (1985) assessed 3-, 4- and 7-year-old children's understanding of (non)factivity using three different off-line tasks. Their results indicate that children from the age of 4 years correctly differentiate between factives and nonfactive mental state verbs (see also Pérez-Leroux and Schulz (1999) and Moore et al. (1989), for similar results; and Harris (1975) and Léger (2008), for different results). One exception is the verb *believe*: the children performed poorly on *belief*-sentences, only adults mastered them. As observed also in Hopmann and Maratsos

(1978), the 3-year-olds apparently employed a *complement-only strategy*, that is they treated the embedded clause to be true irrespective of the verb in the matrix clause. This strategy yields in correct performances on factives but not on nonfactives.

Studies that used sentence repetition revealed that complement sentence acquisition might be first restricted to single CTVs and prototypical, frequent variants of complement sentences (Dąbrowska et al., 2009; Kidd et al., 2005). Kidd et al. (2005) demonstrated that 3- to 5-year-old English-speaking children who were asked to repeat grammatical complement sentences (e.g., *I say he is talking on the telephone.*) and ungrammatical ones (e.g., *\*I say him jumping over the fence.*) corrected ungrammatical sentences more often when the matrix clause contained a high frequency CTVs (such as *think, bet*) compared to low frequent CTVs (such as *pretend, look at*). The authors conclude that these data support the proposal of Diessel and Tomasello (2001), namely that an early lexically-based complementation knowledge provides the basis for the acquisition of grammatical rules yielding in a gradual development of syntactically complex constructions. Interestingly, Dąbrowska et al. (2009) found that in addition to 4- to 6-year-old children even adults do not perform perfectly when repeating complex long-distance questions that entail a 3rd person subject in the matrix clause (e.g., *What does the funny old man really hope you think?*).

Another line of experimental studies examining children's comprehension of sentential complementation have used off-line question comprehension tasks (English: de Villiers, 1999; de Villiers, 1995; de Villiers & Desjarlais, 1995 (reported in de Villiers, 1999); de Villiers and Pyers, 1997; de Villiers and Pyers, 2002; de Villiers et al., 1990; Thornton and Crain, 1994; German: Schulz and Ludwig, 2008). A study by de Villiers et al. (1990) tested children's comprehension of complex wh-questions like in (4) where a question is embedded into the main question.

(4) How<sub>i</sub> did Kermit ask t<sub>i</sub> who to help \*t<sub>i</sub>?

Here, the long distance movement (\*How-help) is blocked because of the medial wh-word (who). Although children aged between 3;7 and 6;11 years have been found to be sensitive to the barrier-function of the medial wh-argument since they behaved adult-like by not answering the long-distance question (\*How-help), younger children often answered the embedded question (\*Who-help) instead of answering the short-distance question in the matrix clause (How-ask), especially when the medial question was an argument (e.g., what, who). Children aged between 3 and 4 years have also difficulties answering long distance questions in which the embedded clause expresses a mistake or a lie as in (5) which also blocks the long-distance interpretation (de Villiers & Desjarlais, 1995 reported in de Villiers, 1999; de Villiers

and Pyers, 1997; de Villiers, Sherrand, & Fretwell, 1992 reported in de Villiers, 1999).

- (5) What<sub>i</sub> did she say t<sub>i</sub> she bought \*t<sub>i</sub>?

In this study the character associated with the subject of the sentence bought something different from what she said. Children aged between 3 and 4 years gave rather an answer to the embedded question (\*What-buy, long distance interpretation) than to the correct short distance question (What-say). As shown by data from de Villiers and Pyers (1997, 2002), the non-adult like performance of the children in this task is obviously not only a matter of the syntactic complexity of the question (see also de Villiers, 1999 for a review). In their *memory-for-complements task* (see 3.1.1) de Villiers and Pyers (2002) asked in addition to the complex wh-question (like in 5) just the simple question "what did she say?" after telling a story like the following example (6):

- (6) (Point at first picture) She said she found a monster under her chair, but (point at second picture) it was really the neighbor's dog.  
(from de Villiers & Pyers, 2002, p. 1043)

Results revealed that only a quarter of 28 normally developing children aged 3;5 years (range 3;1-3;10 years) passed the task, that is they correctly answered either *a monster*, or alternatively *that she found a monster under her chair*; 75 % of the participants still answered incorrectly *a dog* or *that she found the neighbor's dog under her chair*<sup>5</sup>. Around half of the children aged 3;9 and about 90% of the children aged 4;0 solved this task. The authors reported no differences between the two types of question (complex vs. simple question) which supports their idea that it is not an issue of processing limitation but rather a question of whether the assumed PoV feature in the child's grammar is set or not. In addition, in this study children showed no difference between their performances in answering questions containing communicative verbs (*tell* and *say*) and mental state verbs (*think* and *know*). That indicates that it is not about the semantics of the verbs per se but rather the representation of false propositions, which makes the task difficult for children.

Schulz and Ludwig (2008) tested 15 German-speaking children with a translated version of the memory-for-complement task. They presented eight sentences all containing the communication verb *say*. 80% of the children aged 4;2 years (range 3;5 and 4;10) passed the task and thus demonstrated sentential complement under-

---

<sup>5</sup>The criterion for passing was set at 10 or more correct answers out of 12 items.

standing<sup>6</sup>.

To sum up, these findings suggest that children aged between 3 and 4 years old fail to embed the complement under the scope of the matrix verb. They seem to be limited to utterances about real state of affairs since they do not yet fully understand the syntax of complement sentences with its potential to express false propositions.

However, Thornton and Crain (1994, Exp.3) provided evidence that already 3-year-old children comprehend complex wh-questions like (5) correctly. In their experiment, 3;0- to 4;1-year-olds played a "guessing game" with Kermit, a hand puppet. First, the child was asked to hide one of two objects (e.g., a bear). Then, Kermit showed up and was asked to guess which object was hidden. The puppet always guessed something wrong (e.g., a baby). Then the participant was required to answer the test question *What did he say's under there?*. All of the 15 children answered correctly (i.e. *the baby*). The authors concluded that children by age of 3;0 have the syntactic knowledge of complementation. Yet, this interpretation of the data has been criticized by Dąbrowska et al. (2009) who claim that the experimental setting could have led the children to answer correctly because *the baby* was the only new information in the discourse.

In sum, children produce complement sentences early in their language development, but these complement clauses appear to be embedded only on the surface. A full perceptive acquisition of this syntactic structure has been observed only for children around the age of 4 years, slightly before children start to pass standard ToM tasks.

---

<sup>6</sup>The criterion for passing was set at 7 correct answers out of 8 items.





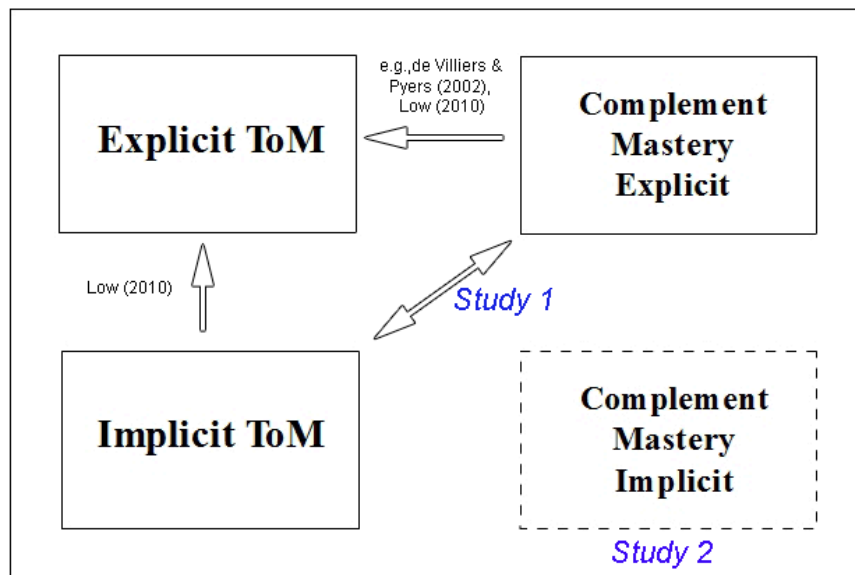
## Chapter 5

# Research Questions and Overview of the Studies

The primary aim of the current work is to investigate if children's implicit ToM is linked to their comprehension of sentential complements. As outlined in Chapter 2, there is an ongoing debate in the ToM literature about the underlying mechanism of children's early success in predictive looking tasks in which no conscious decision is required. A positive correlation between predictive looking and the comprehension of complement sentences would support the hypothesis that a specific syntactic understanding provides the critical means for representing FBs in *any* kind of FB task. In contrast, a lack of correlation between both domains would point towards distinct ToM systems: an explicit one that is language-dependent and an implicit one that might be based on nonlinguistic processes.

Throughout the thesis, 3- to 4-year-olds have been tested because previous research indicates that children in this age range are in a transitional developmental stage as most of them are not proficient in passing explicit ToM tasks yet but display implicit FB understanding. As noted in the introduction, a correlational approach (Study 1) and a methodological approach (Study 2) were applied which are explained in the following. Figure 5.1 displays how the two studies can be considered within the context of previous research.

In *Study 1* (chapter 6), both children's understanding of complement sentences and their implicit and explicit ToM have been collected and tested for interdependence. A new nonverbal predictive looking task was designed to measure participants' anticipatory eye movements. This task differs from previous used implicit FB tasks as it is both completely nonverbal and as complex as the traditional change-of-location task. First, it is hypothesized that children show early FB sensitivity in their eye gaze and thus earlier findings on implicit ToM can be replicated (Clements and Perner, 1994; Low, 2010; Ruffman et al., 2001a; Southgate et al., 2007). Second,



**Figure 5.1:** Overview of the present studies within the context of previous research. The dashed line around the lower right box indicates that there is no study concerning implicit measures of sentential complementation up to now.

in accordance to previous results, a positive correlation between the explicit ToM measure and complement mastery is expected (see connection between the two upper boxes in Fig. 5.1). Concerning the relationship between implicit ToM and sentence comprehension, no specific hypothesis is stated.

Another issue addressed within the scope of *Study 1* concerns the question of which property of sentential complements promotes FB understanding. Previous studies on the developmental link between explicit ToM and complex syntax employed complement tasks that involve misrepresentation (e.g., the memory-for-complement task, de Villiers and Pyers, 2002). That is, the child is required to extract the embedded proposition of a statement or thought that is false with respect to reality (Low, 2010; de Villiers and Pyers, 2002; Hale and Tager-Flusberg, 2003). As has been discussed by Hale and Tager-Flusberg (2003) and recently by Low (2010), it is also plausible that language and ToM are correlated because both measures reflect the ability to handle mistaken propositions (see also Ruffman et al., 2003, and Adler, 2002, for a similar argumentation). This calls the specific impact of complex syntax into question. Therefore, two new complement tasks were developed in order to tap children's knowledge of sentential complementation without relying on misrepresentation (see material for details). A positive correlation between explicit ToM and both tasks is expected supporting the assumption that it is purely the

syntactic device which plays a crucial role for FB understanding.

Moreover, a modified version of the memory-for-complement task was created in which both true and false complements were included to examine whether it is the syntactic structure which enables children to think and reason about FB or whether it is specifically the potency to express false propositions in the complement clause (similar to Tager-Flusberg and Joseph, 2005). According to de Villier's (2005a) proposed developmental progress of complement-taking verbs, it is hypothesized that 3-year-olds perform better on reporting true complements than on reporting false complements. Second, if solely the comprehension of falsity promotes FB understanding children are expected to show better FB task performance when they pass the false complement subtest compared to children who pass only the true complement subtest.

*Study 2* (chapter 7) aims to measure implicit comprehension of complement sentences with the anticipatory eye movement paradigm (first introduced by McMurray and Aslin, 2004). The question addressed here is whether 3-year-olds show any precocious understanding of this complex linguistic structure when tested with the same method as used in predictive looking tasks tapping implicit ToM (see lower right box in Fig. 5.1). Since children's visual orienting reveals a preceding implicit FB understanding it is worth testing whether also an earlier comprehension of sentential complementation can be assessed with indirect measures. To date, there is no available study comparing implicit FB understanding and complement mastery within the same experimental setup. A first step to fill this methodological gap is thus the development of implicit measures of sentential complement comprehension. It is hypothesized that the use of eye-tracking reduces performance demands compared to off-line tasks that rely on answering direct questions, and thus eye-tracking reveals an earlier linguistic competence. That is, analogous to implicit ToM tasks children are not required to make a conscious decision concerning a linguistic representation. In case 3-year-old children show any nascent complement mastery in their eye movements, this opens up a way to expand the syntax-first hypothesis to implicit FB reasoning.

Experiment 1 of *Study 2* tests the scope of the verb *think*. Participants were presented with a complex sentence containing a false proposition and a simple sentence expressing the true state of affairs. After giving the verbal information, it was measured what the participants expected the protagonist to do next. It is hypothesized that children who are proficient in processing the complex sentence in a target-like fashion are guided in their anticipations rather by the content of the embedded proposition than by the the content of the simple sentence. Since the material involves the expression of a mistaken belief, Experiment 1 provides additionally a way to examine whether the implicit ToM sytem is pervious to verbal input. Put differently, it explores whether there is a relationship from the linguistic comprehension to the implicit ToM system that is assumed to control eye gaze but

not children's verbal answers.

In Experiment 2 of *Study 2*, the same method with a new group of participants was used. To control for a potential confound between linguistic and cognitive skills in Experiment 1, Experiment 2 focuses on the scope of negation over a complement clause without giving any information about the truth value of the embedded proposition (e.g., *The girl does not think that the ball is in the red box.*). Thus, children were presented not with two contradicting propositions (i.e. a FB) as in Experiment 1, but instead with one complex sentence expressing a belief. Crucially, including a negative ensures that the processing of the matrix clause (and thus a complement analysis) is necessary for the correct interpretation of the sentence. If children show an increased performance in their visual orienting compared to Experiment 1, it is hypothesized that an additional cognitive load masks children's competence to process complex sentences when tested with complement sentences expressing a false proposition.

## Part II

# EMPIRICAL INVESTIGATIONS



# Chapter 6

## Study 1: Implicit ToM and the Comprehension of Complement Sentences

### 6.1 Introduction/ Aims

The first study of the present work addresses the question whether the correlation found between explicit ToM task performance and the comprehension of sentential complements also holds for implicit FB understanding. Therefore, data on children's comprehension of sentences containing a complement clause and their implicit ToM have been collected and tested for interdependence. As discussed above in Chapter 2, the investigation of a potential relation between language and ToM ability requires a nonverbal FB task. Otherwise high correlations between both measures might be an artefact and reflect rather the fact that the ToM task as such was language demanding.

Assuming that anticipatory eye movements reveal implicit knowledge a paradigm similar to that used by Southgate et al. (2007) and Clements and Perner (1994) was employed: A story's protagonist is distracted while another actor is transferring an object from one container to the other (i.e. unseen-displacement task). Eye gazes were recorded with an eye-tracker while the participant was expecting where the story's protagonist will search for the object after she disappeared behind an occluder. Two cues (a jingling sound and a simultaneous light) that signaled the protagonist's reappearance were included to enhance anticipatory responses. In contrast to Southgate et al.'s (2007) experimental setting, the target object was still in the scene during the anticipation phase. This was done in order to keep the

complexity of traditional ToM tasks. A TB condition was included as a control condition. Here the protagonist witnessed the change of the object's location. Eye movement patterns in the TB scenario can be taken as baseline because the object's real location and the protagonist's belief about the object's location are identical. Any differences in the children's expectations between TB and FB scenarios are attributable to the preceding belief-induction context and thus can be interpreted as evidence of an implicit understanding of belief. Notably, employing this experimental setting entails the possibility to measure not only children's anticipatory looks towards the location they expect the actor to search, but also gazes towards the object's real location (as described in Chapter 2 as *reality bias*). Because these diverge only in the FB condition the paradigm can be seen as a rather conservative test. The first objective of the study, thus, is to investigate preschoolers' performance on this new predictive looking task and to validate whether this experimental setting is suitable for assessing implicit ToM.

Along with the nonverbal implicit ToM task, a traditional verbal FB task was incorporated into the experiment. This was done in order to obtain explicit answers from the children which can be compared with the implicit performance. To rule out any learning effects an unexpected-content task, often referred to as *Smarties task* (Perner et al., 1987), was used instead of the unseen-displacement task.

The assessment of the comprehension of sentential complements contained three different tasks: a truth-value-judgment task, a picture-sentence matching task and a reporting-the-complement task. This allows for methodological comparisons. Since the semantics and the syntax are closely interrelated in mental state verbs, only the communication verb *say* and verb of perception *see* were included in the screening. Thus, it can be ruled out that the comprehension of the meaning of the matrix verb was measured instead of tapping the comprehension of sentential complements.

Another property of verbs that is considered in this study is verb factivity. If the verb in the matrix clause is a nonfactive verb (e.g., *say*, *think*) the embedded clause does not need to have the same truth-value as the main clause. That is, the whole sentence may be true while the complement clause may be false (e.g., *Peter thinks that there is a monster under his bed.*). A modified version of the memory-for-complement task was created (the reporting-the-complement task) in order to determine whether only the comprehension of complex sentences conveying false propositions is crucial or to what extent the comprehension of sentences with true complements is sufficient for predicting FB task performance. According to Tager-Flusberg and Joseph (2005), the former covers both syntactic and semantic features, while the latter is related only to the syntactic property of complement-taking verbs. Both the sentence-picture matching task and the truth-value-judgment task require the child to understand the syntactic structure of the sentences, that is the dependence of the embedded clause from its matrix clause, but these two tasks do not measure the child's ability to understand falsity.



## 6.2 Participants

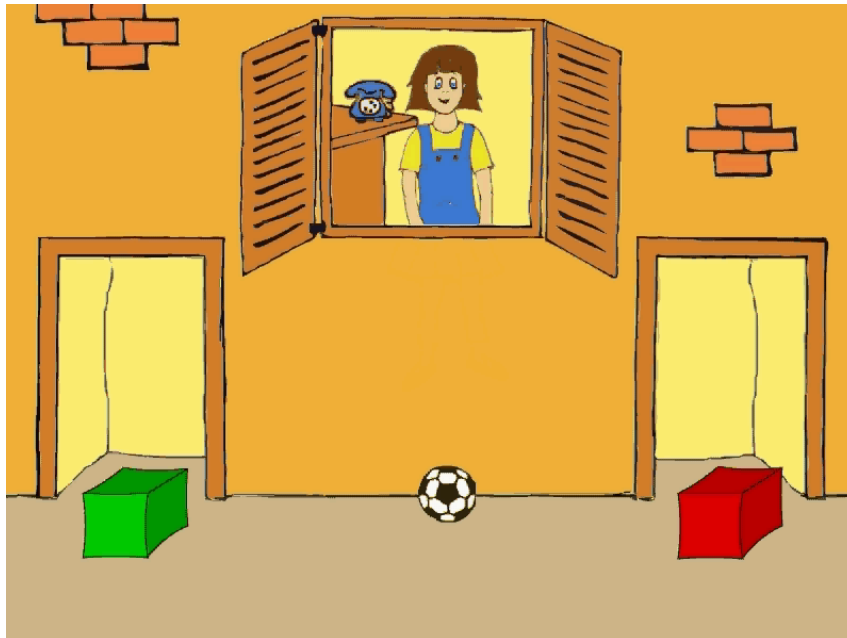
The participants of the first study were 48 monolingual German-acquiring children aged between 3 and 4 years (25 female). The group was split into a younger and an older experimental age group: 25 younger 3-year-olds (mean age: 3 years 3 months (3;3); range: 3;0 - 3;6, 11 female) and 23 older 3-year-olds (mean age: 3;9; range: 3;7 - 4;0, 14 female). In order to validate the new method 29 5-year-old children (mean age: 5;9; range: 5;4 - 5;11, 14 female) and 28 adults (mean age: 23 years, age range: 20-30 yrs, 24 female) took part in this study. Due to different data collection periods only 21 out of the 29 5-year olds (mean age: 5;9, range: 5;8 - 5;11, 10 female) were assessed with the general language tests (described below). All participants were recruited in the Berlin/ Potsdam area.

## 6.3 Material & Procedure for Single Tasks

### 6.3.1 Nonverbal Implicit ToM Task: Unseen Displacement

The nonverbal implicit ToM task was a modification of the *Sally-Anne-Task* (Baron-Cohen et al., 1985) presented as a story in an animated cartoon. There were three different types of items: the familiarization phase without displacing the object, a FB scenario with an unseen object displacement and a TB scenario with a seen object displacement. All parts were completely nonverbal, that is without voice-over narration and without comments about the protagonist's intention and beliefs. In order to maintain the attention of the participants there were some nonlinguistic background sounds, e.g. steps or the twittering of birds; furthermore, the hidden objects varied over trials (puppet, ball, toy car, toy duck, pacifier, teddy bear). Participants' eye movements were monitored and recorded with an eye-tracker while they were watching the video clips. Implicit, behavioral measures of action anticipation served as core dependent variables. Additionally, like in the original *Sally-Anne-Task* a test question was posed in verbal form in a FB scenario at the end of the experiment (see paragraph *False Belief Condition*).

**Visual Scene** The visual scene presented in the video clips is shown in Figure 6.1. It consists of a house's frontview with a window on the first floor and two exits on the ground floor, each exit with a nearby box in front: a green left-hand box and a red right-hand box. An object is located between the two boxes. The protagonist of the story, a brown-haired girl, is looking out of the window. A telephone is located behind the girl on a table.



**Figure 6.1:** Visual scene of the implicit nonverbal ToM task

**Familiarization** There were two familiarization trials in total (see Figure 6.2 A). They differed in the place the object was hidden. At the beginning of a familiarization trial an object was situated on the ground in front of the house between the two boxes. A blond-haired girl who entered the scene from the bottom picked up the object and hid it in either the red box or the green box (Fig. 6.2 A/1). The hiding procedure was watched by the protagonist. To emphasize that the protagonist was attending to what was happening outside the house, her eyes always followed the movement of the blond-haired girl. Note that both boxes were closed, so that the brown-haired girl could not look inside the box from above. After the blond-haired girl disappeared from the scene, the girl at the window made a noise (“*hmm*”), turned around and started to move downstairs (Fig. 6.2 A/2). While the protagonist was disappearing a chime sounds and both frames were illuminated simultaneously. After a 1000-ms delay the protagonist reappeared in that exit corresponding to the box in which the object was hidden (Fig. 6.2 A/3). The girl opened the box, found the object and was very pleased to obtain it (she laughed and jumped up and down).

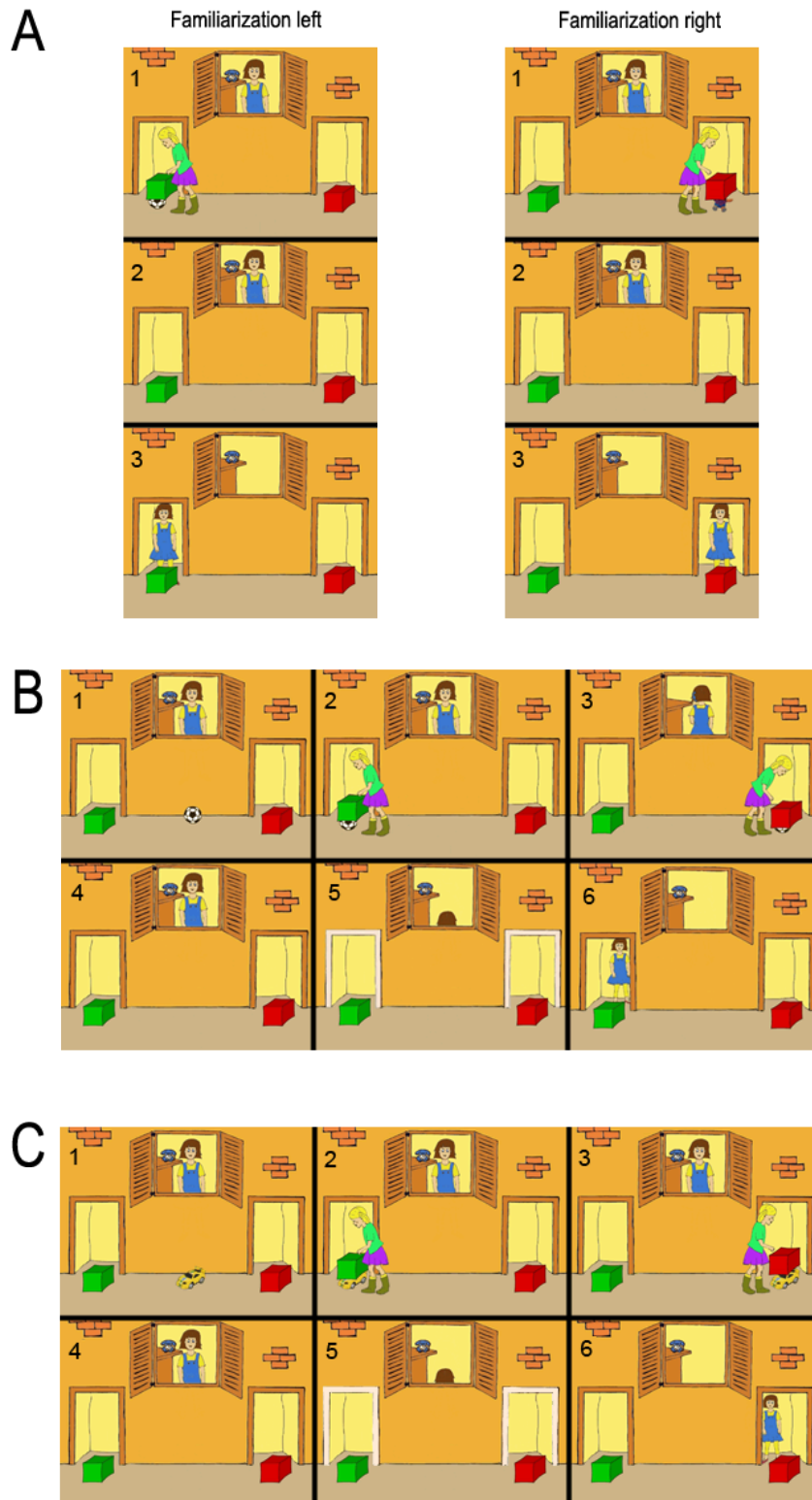
The purpose of the familiarization trials was to teach the participants (a) that the girl in the house wants to obtain the object after she starts to move and (b) that she has access to both exits and always takes the shortest way to that box she believes to find the object in. Moreover, participants were supposed to get familiar with the fact that after the frames were illuminated and the simultaneous cue had occurred the girl was about to show up again in one of the two exits.

**False Belief Condition** In the FB condition (illustrated in Figure 6.2 B) the first hiding sequence was the same as in the familiarization part. After the hider had disappeared, the telephone rang and the girl in the house turned around and talked to someone on the phone mumblingly. While she was distracted the blond-haired girl came back, opened the box in which the object was placed, retrieved the object and placed it into the other box (see Fig. 6.2 B/3). Then, after the hider had left the scene again, the girl in the house stopped phoning, went back and looked outside the window (see Fig. 6.2 B/4). The girl - now holding a FB about the object's location - started to move downstairs and disappeared for 1600 ms. Compared to the familiarization phase the critical time window for anticipation was extended in the test trials (in both the FB and TB condition) in order to enhance anticipatory eye movements.

In the *anticipation-only* FB trial, the clip stopped after showing the place of the girl's reappearance. Hence, the information whether the girl discovered the object was not given. This was done to eliminate training effects. In the *anticipation + prediction* FB trial, the girl did not show up again after she disappeared. This unresolved version of the story was employed to elicit a conscious, verbal reaction from the participants by asking them to predict the side of reappearance. After the 1600-ms anticipation-window has passed a child-directed background voice was prompting a prediction by questioning: *Zeig mir mal, wo kommt das Mädchen jetzt raus?* ('Show me! Where will the girl show up again?'). Note that the anticipation phase was identical in both FB versions and that the version including the additional prediction was always presented at the end of the implicit ToM task to ensure that demanding a conscious, explicit answer does not interfere with eye movement behavior during the experiment<sup>1</sup>. Again, there were two realizations of this condition: one trial in which the place of reappearance (i.e. the first-hiding place) is the left-hand box (as in the example in Fig. 6.2 B), and one trial in which it is the right-hand box.

---

<sup>1</sup>Brandt-Kobe and Höhle (2010) provided evidence that an additional explicit task interferes with childrens looking behavior.



**Figure 6.2:** Selected frames from the events of the implicit ToM task shown in the (A) familiarization phase, (B) FB condition (place of reappearance: left), and (C) TB condition (place of reappearance: right)

**True Belief Condition** The TB condition is identical to the FB condition except that the girl in the house was not distracted by a ringing telephone but attended to the scene while the object was displaced by the other girl (see Figure 6.2 C). The place of reappearance is thus identical with the last-hiding place. As in the FB condition (*anticipation-only* version) and again in order to avoid possible training effects, the movie stopped after the girl had reappeared, not showing whether the girl succeeded in finding the object. Two versions of the TB scenario were constructed, a left-hand and a right-hand version (see the right-hand version exemplified in Fig. 6.2 C).

Each participant received two familiarization trials, two FB trials and two TB trials; the place of reappearance was the left and right door in each case. The order of the testing trials was pseudorandomized, i.e. the last test trial was always a *anticipation + prediction* FB trial and there were no two FB trials presented consecutively, yielding eight possible orders.

### 6.3.2 Explicit Verbal ToM Task: Unexpected Content

As a verbal explicit FB task an *unexpected-content task* was employed: the *Smarties task* (Perner et al., 1987). As in the original task, a Smarties tube was shown to the participants and the children were then asked what they think is in the box. After the child had usually answered “Smarties” or “sweets” it was shown that the tube in fact contained pencils. The box was reclosed then and the experimenter asked the child (a) what s/he had thought was inside the box before it was opened (Smarties *self*) and (b) what another person (e.g. the mother or father) who had not yet seen the tube’s content will think what is inside the box (Smarties *other*). The German translations of the original test questions were the following: *Was hast Du gedacht, ist hier drin, als ich Dir die Schachtel gezeigt habe?* (*self*); *Was glaubt die Mama/ der Papa ist hier drin, wenn wir ihr/ihm die Schachtel zeigen?* (*other*). A child passed the task if s/he answered “Smarties” or something comparable like “sweets” indicating that s/he ascribed a (former) false belief to her/himself and/ or to another person. Instead, children demonstrated a lack of FB understanding when they responded with the real content of the box: “pencils”.

### 6.3.3 Complementation Tasks

Three language tasks were conducted to assess children’s comprehension of tensed sentential complements. Each task applied a different method: truth-value-judgment, picture-sentence matching and reporting-the-complement. Two of them (truth-value-judgement task, picture-sentence matching task) required the comprehension of the *syntactic* property of complement constructions, that is that a complement clause functions as an argument of a matrix verb and thus the embedded proposition

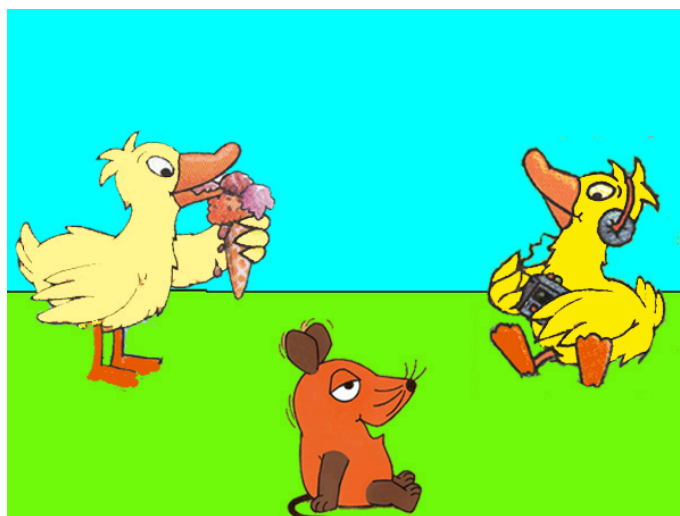
can not be interpreted as independent sentence. The third task (reporting-the-complement task) is a modification of the memory-for-complement task and assesses in addition to the syntactic knowledge also children's understanding of a *semantic* property of complement clauses of nonfactive verbs: the embedded clause may have an independent truth value, that is the proposition of the embedded clause may be false while the whole sentence is true. Both true and false complements were included in this task in order to examine the effect of both properties of sentential complements. In addition, this task also includes a productive part since it requires the child to report a complement. No mental state verbs were incorporated into the tasks in order to avoid interference with the semantic-conceptual comprehension of cognition verbs and thus with ToM abilities. Instead, a verb of communication (*say*) was used in the reporting-the-complement task and a verb of perception (*see*) in the two other tasks.

### 6.3.3.1 Truth-Value-Judgment Task

In total, 12 pictures were shown consecutively, one at a time. As exemplified in Figure 6.3, every drawing displayed three actors: two animals of the same kind, both are doing something different at the top of the picture (e.g., the ducks) and another animal at the bottom of the picture (e.g., the mouse) that is looking at one of the other characters. Each picture was paired with a sentence containing a sentential complement under the perception verb *sehen* (*to see*). In half of the items, the sentence expressed the situation on the picture correctly (like in 1a), the other half did not match with the picture (1b):

- (1) a. *Die Maus sieht gerade, dass die Ente Musik hört.*  
 'The mouse is seeing that the duck is listening to music.'  
 b. *Die Maus sieht gerade, dass die Ente Eis isst.*  
 'The mouse is seeing that the duck is eating ice cream.'

Participants were asked to judge whether the sentence matched with the presented picture or not. Six items required a *yes*-answer, the other half required a *no*-answer. The proposition of the embedded clause was depicted on each drawing. To judge the sentence correctly, the child has to interpret the embedded clause as the object of the matrix clause. In case of a high cognitive load, children may not parse the complex sentence, but instead attend only to the embedded proposition and thus rather give *yes*-answers which would lead to correct responses in the matching-condition but not in the non-matching condition. Each picture was used in a *yes*- and in a *no*-trial for each participant to rule out material-specific responses. Additionally, the animal associated with the subject of the sentence (the mouse in the example) was depicted



**Figure 6.3:** Example of a picture presented in the truth-value-judgment task.

reversely so that it gazed towards the second animal (ice-eating duck in the example). This time, consequently, the former correct description was now incorrect and vice versa. In sum, there are three scenarios in total, each occurring with two types of visual orientation of the sentence's subject referent (left/right) and two types of test sentence (matching/non-matching). The complete list of all 12 items is given in Appendix A.1.

In order to encourage children to reject a sentence, they were introduced to a hand puppet called Erwin, who does not speak well yet because he is still learning German. The child was asked to help Erwin by rewarding him with a wooden banana whenever his description of the picture was correct and by punishing him with a wooden tomato in case he made a false statement about the picture (Gordon, 1996, reward/punishment version). The test sentence was presented by the experimenter while moving the hand puppet's mouth. The experimenter then asked the child: *Was gibst Du dem Erwin jetzt? Die Tomate oder die Banane?* ('What are you giving to Erwin now? The tomato or the banana?'). Children did not need to answer verbally; in this case the chosen food was taken as yes/no response. For the test items, no feedback was given. Note that every participant received all possible picture x sentence combinations (12 items in total). Two versions were constructed by varying the order of the presented pictures: if in version A/Item 1 the sentence matched with a picture (correct *yes*-answer), the same sentence did not match in version B/Item 1 (correct *no*-answer) (the order for each version is given in Appendix A.1). Half of the subjects were assigned to version A; the other half to version B.

One introductory picture and four practice trials were presented at the beginning

of the task (also listed in Appendix A.1). Children were asked to point to the labeled animals in the introductory picture. This was done in order to ensure that the children knew the animals (ducks, elephants and mice) and to familiarize them with the fact that there were two instances of each kind. In the practice trials, participants were shown pictures displaying the three animals (e.g., the duck and the mouse that is looking at the elephant) and simple SVO-sentences that had to be judged concerning their correctness (e.g. *Die Maus sieht gerade den Elefanten.*/ ‘The mouse is seeing the elephant.’). In doing so, it was ensured that the subjects interpreted the viewing direction of the animal at the bottom of the picture correctly. Two out of four descriptions were correct, two were incorrect; the participants received feedback in the practice trials. The pictures were presented on cards; sentences were spoken by the experimenter.

### 6.3.3.2 Sentence-Picture Matching Task

This task was a modification of a subtest from the TROG-D (Test for Reception of Grammar, German version, Subtest U Subordination; Fox, 2006) and consisted of four items. Participants were shown three pictures per item (as illustrated in Figure 6.4), all portraying the same persons (in the example: Mommy and Lisa), but differing concerning the orientation of the persons and the agent-patient relation, respectively. The task was to select that picture which best matched with a test sentence containing a sentential complement under the verb *sehen* (*to see*) like in (2).

- (2) *Die Mama sieht, dass Lisa auf sie zeigt.*  
‘Mommy sees that Lisa is pointing at her.’

There were two types of distractors: one picture illustrated the embedded clause correctly, but not the complete sentence (in the example the lower picture). The other distractor can not be characterized in a homogeneous fashion. In the example, the action of the embedded clause was reversed (the upper right picture). A list with all items in their presented order appears in Appendix A.2. In order to introduce the persons on the pictures (Lisa, Tim, Mommy), an introductory picture was presented at the beginning of the task. All visual and auditory stimuli were presented on a laptop using PowerPoint. To motivate the children the same hand puppet (Erwin) as in the truth-value-judgment task was used; this time the children got the information that he has “some problems with his eyes and needs some help in finding the proper picture”.



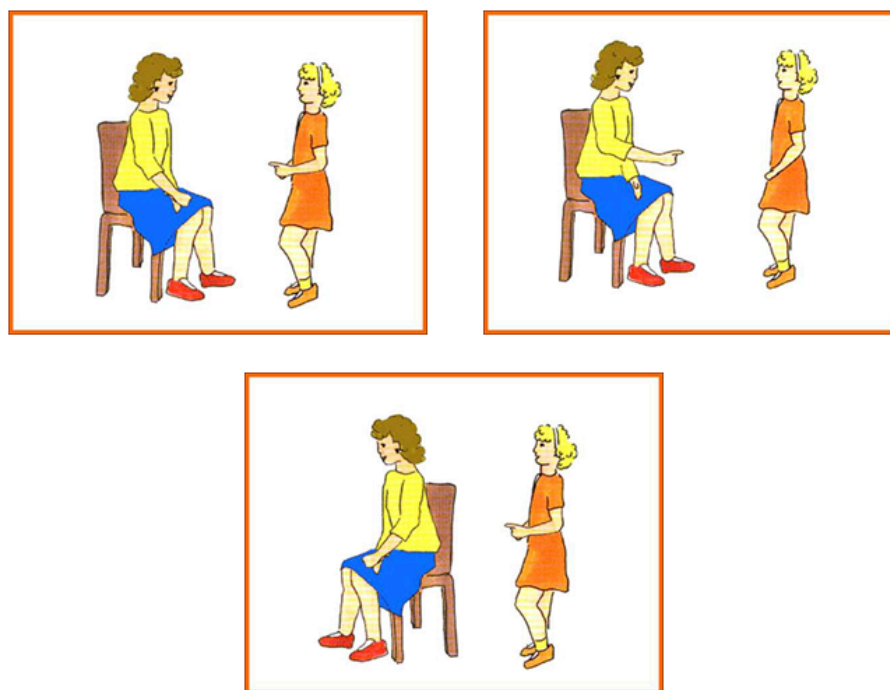


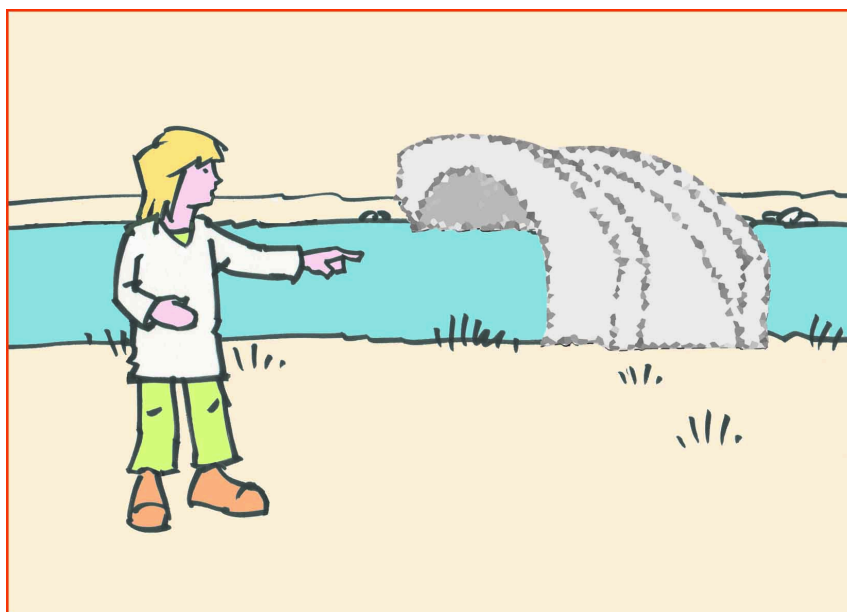
Figure 6.4: Example of an item presented in the sentence-picture matching task

### 6.3.3.3 Reporting-the-Complement Task

This task is based on the memory-for-complements task developed by Jill de Villiers and colleagues (de Villiers and de Villiers, 2000; de Villiers and Pyers, 2002). Children were presented with a total of eight brief stories in which the protagonists are described as making a statement about a situation like in (3 a). Critically, a simple sentence like in (3 b) about the real state of affair was presented, too. Each story was accompanied by a drawn picture as shown in Figure 6.5 for (3).

- (3) Intro *Das ist Julia.*  
 ‘This is Julia.’
- a *Julia sagte, dass die Brücke kaputt ist.*  
 ‘Julia said that the bridge is broken.’
- b *Die Brücke war in Wirklichkeit ganz.*  
 ‘Really the bridge was undamaged.’

The child then was asked to report the complement: *What did [name of the story’s protagonist] say?* Again, the hand puppet Erwin was used for motivation purposes; this time he stated that he is “hearing impaired and does not get all the information



**Figure 6.5:** Example of an item presented in the reporting-the-complement task

of the stories”.

Originally, the task requires the children to remember and extract a false embedded proposition from sentences containing mental state verbs and verbs of communication. In contrast to the original version, only the communication verb *say* was used here to avoid a confounding with the child’s ability to handle the semantics of mental state verbs. As a further modification, the pictures illustrated the crucial issue ambiguously, i.e. it was not apparent whether the bridge in example (3) was broken or not. This was done to avoid reality intrusions caused by the depiction of the true state of affairs.

Furthermore, and similar to Tager-Flusberg and Joseph (2005), both true (i.e. consistent with the actual world), and false complements (i.e. a lie or mistake) were included. This allows to parcel out the semantic and syntactic properties of sentential complements. The false-complement condition tapped both semantic and syntactic knowledge since the child had to extract the content of the complement which was incompatible with reality - this condition is called *2-representation-condition*. The true-complement condition (or *1-representation-condition*) tapped syntactic knowledge of complementation since there was no conflict between the reality and the content of the embedded clause. In the study of Tager-Flusberg and Joseph (2005), the embedded (true) clause had almost the same wording as the simple sentence describing the true state of affairs (e.g. Jeff said, “*The lights are broken.*”. The lights were really broken. What did Jeff say?). Thus, it is not evident whether

the child's response referred to the simple sentence or the embedded complement. To ensure that children really extracted the true complement from the complex sentence, the sentence describing reality (4b) paraphrased the content of the embedded clause (like in 4a):

- (4) a. *Peter sagte, dass die Sonne scheint.*  
'Peter said that the sun is shining.'  
b. *Draußen war wirklich schönes Wetter.*  
'The weather was really nice outside.'

The order of the complex sentence containing the protagonist's statement and the simple sentence describing the true state of affairs were counterbalanced across items within participants to avoid that children simply report the last sentence. Two versions of this task were constructed: when in version A /item X the simple sentence preceded the complex sentence, the order was inversed in version B, and vice versa. Like in Tager-Flusberg and Joseph (2005), the word "really" was used in the simple sentences; it was translated into "in Wirklichkeit" for items of the 2-representation-condition and into "wirklich" for items of the 1-representation-condition. The complete list of items is given in Appendix A.3. The items were presented in a fixed order, and the first two items were warm-up items (one item from the 1-representation-condition and one from the 2-representation-condition) in which no feedback was given. Six test items followed. All visual and auditory stimuli were presented on a laptop using PowerPoint. A response was scored correct in case the child reported the sentential complement (indirect or direct quote) or just a part of the sentential complement, e.g. *Sonne scheint* ('shining sun') for story (4). Thus, the number of correct responses could range from 0 to 6; for each condition from 0 to 3.

### 6.3.4 General Language Measures

Two standardized language tests were administered: the TROG-D to assess the understanding of grammatical structures in German (Fox, 2006) and the AWST-R for the assessment of the productive vocabulary (Kiese-Himmel, 2005). These scores served later as variables to validate the sentential complement comprehension tasks and additionally as a control to rule out delayed language acquisition as a potential factor. The parents received a test evaluation by letter.

#### 6.3.4.1 Test for Reception of Grammar: TROG-D

Comprehension of grammar was measured by the TROG-D (Test for Reception of Grammar, German version; Fox, 2006). The children were presented with a sentence

spoken by the experimenter and an array of four pictures and then asked to select that picture that matches the sentence best. Since the child can respond by pointing, no expressive language is needed in this test. The three distractor pictures differ from the target picture with regard to grammatical relations or lexical content. In 21 subtests the understanding of different aspects of grammar like word order, relative clauses, pronouns, inflections, embedding etc. are assessed; the test has norms for an age range of 3;0 to 10;11 years.

#### 6.3.4.2 Active Vocabulary Test: AWST-R

The AWST-R (Aktiver-Wortschatz-Test Revision; Kiese-Himmel, 2005) was used to assess children's active vocabulary. This instrument is standardized for 3;0 to 5;5 year-old children and consists of photographs depicting 51 nouns and 24 verbs. Participants were asked to label the illustrated noun or action. Since the children of the control group were sometimes older than 5;5 years, the validity of the test was constrained in these cases.

## 6.4 General Procedure and Apparatus

All participants were tested individually in a testing room located at the University. Most of the time the children's parents were present while testing, though they were asked not to influence the child's reactions. For children of both the experimental and the control group, three or four sessions were required to complete the testing, depending on the concentration of the child. It was intended that a maximum of 7 days passed between two successive sessions. This was not possible in all cases due to sickness or other unpredictable circumstances.

Within the first session, the eye-tracking experiment and optionally the TROG-D were conducted, because both tasks did not require the child to answer verbally. Thus especially shy children got the opportunity to get used to the testing situation and/or the experimenter without necessarily speaking to her. Prior to the eye-tracking experiment, a color recognition task was presented to each participant to ensure that the looking behavior does not result from lack of color knowledge or red-green color blindness. Here, participants had to point to a red and a green elephant on a drawing. At the end of the first session the parents filled in a questionnaire about the child's development. In the second session the language tests assessing children's understanding of sentential complements and the Smarties task were conducted. In the third session the AWST-R was administered; if necessary, there was a fourth session for the completion of the TROG-D. Adults were tested in one session; they were not tested in the language tests and the Smarties task.

During the eye-tracking experiment, each participant was seated in a leaned-back

chair, 60 cm in front of a monitor on which the nonverbal ToM movies were presented. Children were instructed to simply watch the cartoon. To collect the eye gaze data from the children, a Tobii 1750 Eye Tracker (Stockholm, Sweden) was used, which is integrated into a 17-in. monitor and provides eye gaze position data 50 times per second. Stimuli were presented in avi-format with the Clearview 2.5.1 software package, which was used for data collection as well. For adults, a Tobii T120 Eye Tracker along with the Tobii Studio software was used for stimuli presentation and eye gaze recording. The Tobii T120 Eye Tracker tracks the position of the eyes 120 times per second so that the position of the left and right eye is provided every 8.3 milliseconds. Before starting the experiment, a five-point-calibration was performed for all participants. Between trials, animated cartoon figures were presented for recapturing or increasing children's attention and to mark the beginning and ending of the experimental trials. The entire eye-tracking experiment took about 12 minutes.

During the language tasks, the child participants sat at a table suitable for children. In the second session, each child conducted first the truth-value-judgment task, followed by the reporting-the-complement task and finally the sentence-picture matching task. It took about 25-35 minutes to complete the complementation tasks. Children's responses were recorded at the time of testing. In order to create a realistic situation for the Smarties task the parent had been instructed to leave the room for plausible reasons (forgot something, go to the bathroom etc.). Meanwhile, the child was asked to predict what the parent will think is inside the box. The second session was videotaped.

## 6.5 Data Analyses

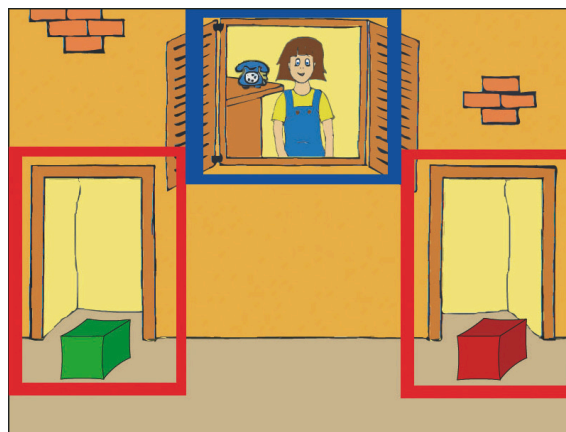
In a first step, performance on each single task was analyzed separately to obtain an overview of the data. Anticipatory eye gaze data of the nonverbal implicit ToM task were aggregated and analyzed using the R software (R Development Core Team, 2009). Overall, the data were analyzed in a number of ways: group differences were explored with ANOVA, Chi-square test and Mann-Whitney-U test. One-sample Wilcoxon test and binomial test were conducted to determine whether performance differed from chance level. Correlations were computed between language measures and ToM tasks. To estimate the relation between the sentential complement tasks and the explicit ToM task within the 3-year-olds, a hierarchical logistic regression analysis (with age as control variable) was conducted using SPSS. Due to the small sample size, the logistic regression was limited to two predictors: age and a composite score of the sentential complement tasks.

## 6.6 Results

### 6.6.1 ToM Measures

#### 6.6.1.1 Nonverbal Implicit FB Task

**Data Reduction** In order to analyze *where* and *for how long* the participants anticipated the reappearance of the girl, two spatial areas of interests (AOI) were defined (250 x 350 Pixel), one around each exit including the corresponding box (framed in red in Fig. 6.6). One of the AOIs is referred to as *first-hiding AOI* and the other one as *last-hiding AOI*. The correctly anticipated place of reappearance was dependent on the condition, that is the last-hiding AOI was the correct AOI for the TB context and the first-hiding AOI for the FB context. Additionally, an AOI around the window was defined (framed in blue in Fig. 6.6) in order to check whether the participants followed the story and the critical information whether the girl upstairs is watching the replacement or not. The temporal window was defined to be 1600 ms, i.e. the time period after the girl disappeared from the window until she reappeared at one of the exits. Two core measures of action anticipation were calculated for each trial and participant: the direction of the first look (correct vs incorrect) and the amount of time the participant spent on each of both AOIs. Here, fixations were detected by the software packages and the proportions of fixations to the first- and last-hiding AOIs were calculated<sup>2</sup>. The criterion for a fixation was set to a minimum duration of 60 ms and a radius of 50 pixel.



**Figure 6.6:** Areas of Interest (AOI): first-hiding place and last-hiding place framed in red and the control AOI around the window framed in blue.

<sup>2</sup>All gazes towards a certain AOI were relativized to all *valid* gazes on screen within this time span.

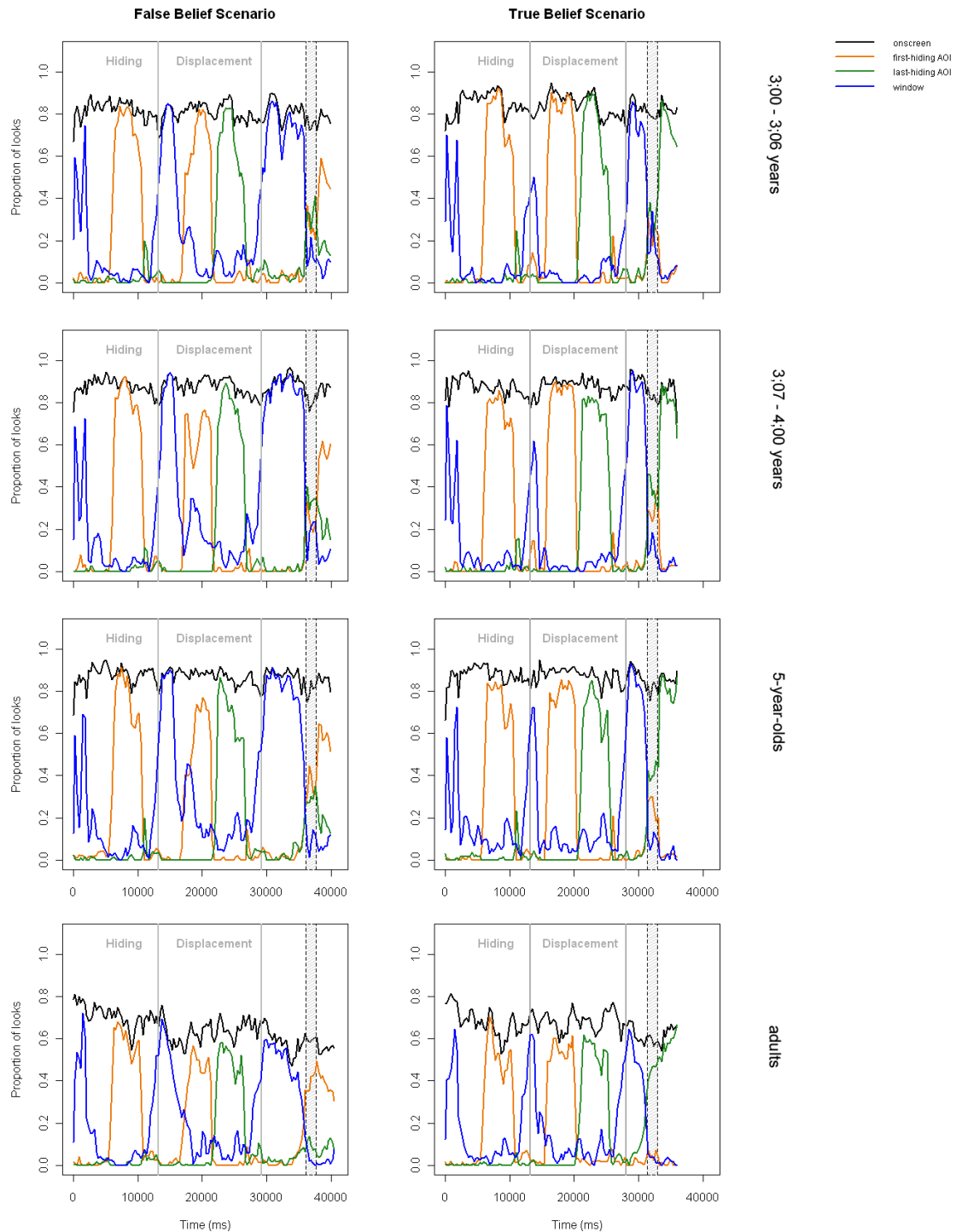
Since the looking proportions to both AOIs within the same trial are not independent of each other and since the TB condition is considered as baseline, the difference of the looking proportion towards one AOI between both belief conditions was analyzed. If children possess an implicit ToM they were expected to anticipate the reappearance of the protagonist to a higher amount in the first-hiding AOI in the FB condition compared to the TB condition. Simultaneously, looking proportions to the last-hiding AOI are expected to be greater in the TB condition than in the FB condition. Thus, it was hypothesized that there is an interaction effect between condition and AOI, but no main effects of condition and AOI.

**Time course** First, a descriptive analysis of the time course of the testing trials was conducted. Therefore, the gazes within each condition, the proportion of fixations to the first-hiding AOI, to the last-hiding AOI and to the window were calculated for 300 ms time slices, aggregating trials for each condition first within a participant and then across participants. Time course plots are given in Fig. 6.7 for the FB condition<sup>3</sup> (left panel) and TB condition (right panel) and each participant group (each row). The light grey area represents the time span in which anticipatory eye movements occur. Participants of all age groups followed the story in both conditions: During the first hiding sequence there was a high amount of looks to the first-hiding AOI, during the displacement sequence the eyes went from the first-hiding to the last-hiding AOI and back to the window. Additionally, looks towards the window during the displacement phase were more likely in the FB condition. Whereas in both experimental groups children's anticipatory looks towards the last-hiding AOI tended to be more likely irrespective of belief context, participants of the control groups tended to expect the protagonist to reappear in that AOI where she last saw the object being placed. Interestingly, adults showed correct anticipatory eye movements even before the girl has disappeared completely. This is even more obvious for the TB trials. These early anticipations were not present in the children's data.

**First looks** Table 6.1 gives the numbers and percentages of trials with correct and incorrect first looks for each age group and condition. Note that cases in which the number of correct and incorrect first looks do not add up to the original number of trials within each age group are due to participants who did not anticipate at all. Two-choice binomial test (two-tailed) were conducted to test for significance. None of the child participant groups showed a difference between correct and incorrect first looks in both the FB- and TB condition (all  $p$ 's > .1). Adults gazed more often toward the belief-accordant AOI in both conditions (FB:  $p < .001$ ; TB:  $p < .001$ ).

---

<sup>3</sup>Note that both the *anticipation + prediction* and the *anticipation only* FB trials were averaged causing no clear correct looking behavior after the anticipation phase ended.



**Figure 6.7:** Average looking proportion during a course of false-belief trials (left) and a true-belief trials (right), separate for younger 3-year-olds (upper row), older 3-year-olds (second row), 5-year-olds (third row) and adults (lower row). The black line indicate the proportion of valid fixations, the orange line indicates the proportion of all fixations within the first-hiding AOI, the green line for gazes towards the last-hiding AOI. Gazes towards the window are indicated by the blue line.



**Table 6.1:** Number of trials with correct and incorrect first looks (percentages in brackets), split by age group and condition. *Note.* \*\*\* $p \leq .001$ .

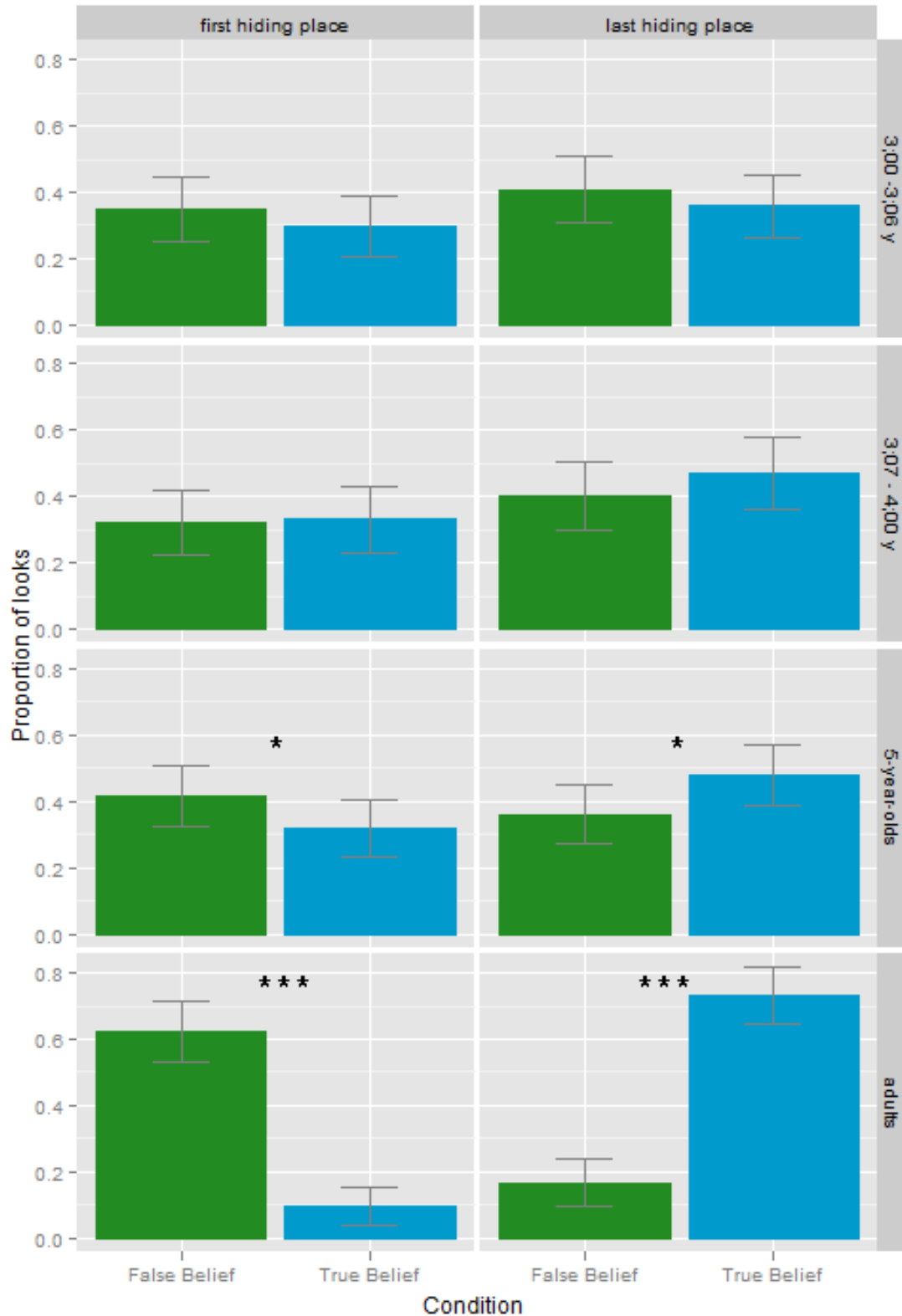
		3;00-3;06 y (50 trials)	3;07-4;00 y (46 trials)	5-years old (58 trials)	adults (56 trials)
False Belief	correct	27 (61%)	19 (44%)	28 (49%)	<b>35 (80%)***</b>
	incorrect	17 (39%)	24 (56%)	29 (51%)	9 (20%)
True Belief	correct	21 (45%)	23 (52%)	31 (55%)	<b>37 (86%)***</b>
	incorrect	26 (55%)	21 (48%)	25 (45%)	6 (14%)

**Looking proportions during anticipation phase** Out of the 105 participants in total, one adult and one three-year-old girl (3;9 y) were excluded from the analysis because of missing data<sup>4</sup>. Furthermore, valid data from one participant of the younger testing group (female, 3;4 y) was available only for the TB condition, but not for the FB condition. This child, however, was not removed from the analysis. The average fixation proportions at the first- and last-hiding place from each participant group and condition are illustrated in Figure 6.8. First, a mixed-effect three-way analysis of variance (ANOVA) with group (3;0y–3;6y; 3;7y–4;0 y, 5-year-olds, adults) as between-subject factor and AOI (first- vs. last-hiding place) and condition (FB vs. TB) as within-subject factors was applied to this data using the R software (R Development Core Team, 2009).

No main effect of *age group* was obtained ( $F(3, 98) = 1.49; p > .1$ ), i.e. all age groups spent approximately the same amount of time (proportionally speaking) within the AOIs during the anticipation phase (3;0–3;6 y: .35; 3;7–4;0 y: .38; 5-year-olds: .39; adults: .40). There was also no significant main effect for *condition*:  $F(1, 98) < 1$ , i.e. looking proportions in both the FB condition and the TB condition were similar (in each condition: .38). The interaction of *condition* with *age group* was marginally significant ( $F(3, 98) = 2.41; p = .07$ ). Both adults and 5-year-olds displayed similar anticipation proportions in the FB and TB condition (adults<sub>FB</sub>: .39; adults<sub>TB</sub>: .41, 5-year-olds<sub>TB</sub>: .39; 5-year-olds<sub>FB</sub>: .40), whereas the younger testing group had greater anticipation proportion in the FB condition (TB: .33; FB: .38) and older 3-year-olds displayed the reverse pattern (TB: .40; FB: .36).

Across all age groups and conditions, there was a significant main effect of *AOI*, i.e. a difference between the looking proportions towards the first-hiding place (.34) and the last-hiding place (.42) ( $F(1, 98) = 9.36; p < 0.01$ ). There was no significant interaction between *AOI* and *age group*, i.e. all age groups showed higher looking proportion towards the last-hiding place ( $F(3, 98) < 1$ ). The interaction between *AOI* and *condition* turned out to be highly significant ( $F(1, 98) = 30.65; p < .001$ ), i.e. in the FB condition there were higher looking proportions for the first-hiding

<sup>4</sup>This is due to technical problems.



**Figure 6.8:** Average looking proportion with reference to the total looking time towards first and last hiding place in false-belief trials (green panels) and true-belief trials (blue panels); separate for younger 3-year-olds (upper row), older 3-year-olds (second row), 5-year-olds (third row) and adults (lower row); error bars indicate standard errors for proportions. *Note.*  $*p \leq .05$ ,  $***p \leq .001$ .

place (.43) compared to the last-hiding place (.33), whereas for the TB condition there were higher looking proportions for the last-hiding place (.51) compared to the first-hiding place (.26) when averaging across age groups. In addition, the three-way interaction *AOI* x *condition* x *group* was found to be highly significant ( $F(3, 98) = 15.48; p < .001$ ).

A 2 x 2 repeated measures ANOVA (condition x *AOI*) for each age group separately yielded a significant interaction effect between *AOI* x *condition* for adults ( $F(1, 26) = 42.02; p < .001$ ), and 5-year-olds ( $F(1, 28) = 4.12; p = .05$ ). This was not found for both testing groups (3;0–3;6 y:  $F(1, 21) < 1$ ; 3;7–4;0 y:  $F(2, 23) < 1$ ). There was no significant main effect of *condition* and *AOI* for either age group (all  $p$ 's  $< .09$ ). Thus, the overall interaction between *AOI* and *condition* was due to the looking data of the adults and 5-year-olds. An additional 2 x 2 ANOVA of collapsed data across 3-year-olds revealed a significant main effect of *AOI* ( $F(1, 45) = 4.19; p = .05$ ), but no main effect for *condition* and no interaction effect (both  $F$ 's (1,45)  $< 1$ ). That is, irrespective of belief condition, 3-year-olds anticipated the protagonist to go to where the object really is (last-hiding-place: .41; first-hiding place: .32).

Follow-up paired t-tests (one-tailed) between the FB-condition and the TB-condition revealed that adults looked significantly longer to the first-hiding place in the FB condition (.62) compared to the TB condition (.10) ( $t(26) = 6.03; p < .001$ ). 5-year olds also show a significant difference in the same direction ( $t(28) = 1.83; p < .05$ ; FB: .41; TB: .32). As for the last-hiding place, both adults and 5-year-olds had significantly higher looking proportions in the TB condition compared to the FB condition (adults:  $t(26) = -6.63; p < .001$ , FB: .17; TB: .73; 5-year-olds:  $t(28) = 2.03; p < .05$ , FB: .36; TB: .48).

**Interim summary of eye-tracking data** This nonverbal eye-tracking task was designed to measure childrens implicit understanding of FB. Results on looking times showed that 5-year-old children and adults anticipated the protagonist to reappear at that location which corresponded to her belief. Interestingly, the correct expectation was reflected in first looks only for adults but not for 5-year-olds. Thus, in general anticipatory eye movements are a sensitive measurement for ToM processing as has also been shown in previous studies (Clements and Perner, 1994; Low, 2010; Ruffman et al., 2001a; Southgate et al., 2007). However, the data did not provide evidence for an early implicit ToM reasoning in 3-year-olds. In both belief conditions, they anticipated the protagonist to reappear at that location where the object really was. That indicates that they were not taking the protagonists (false) belief about the object's location into account.

This result is in conflict with findings of an earlier onset of implicit FB understanding starting around the age of 2;11 (Clements and Perner, 1994) or even earlier around age 2;1 (Southgate et al., 2007). Since both control groups displayed the expected looking pattern in the current study it is unlikely that the task failed to tap ToM reasoning per se. Comparisons between the present study and the ones that found evidence for an early implicit ToM reveal differences between the experimental setting. The result suggests that the characteristics of predictive looking tasks play a role for the elicitation of implicit FB understanding; this will be discussed in section 6.7.

### 6.6.1.2 Explicit FB Tasks

First, analyses of participants' performance will be reported on each task and age group separately. Here, binomial tests were used to examine whether performance differed from chance (50%). Secondly, in order to analyze whether there was a difference in task difficulty between Smarties *other* and Smarties *self*, McNemar tests were performed for each group of children. To compare performances between age groups, chi-square tests (two-tailed) were used.

Table 6.2 gives the number and percentage of participants for each age group who answered the Smarties *self* FB question correctly, the Smarties *other* FB question and the action prediction in the last FB trial of the eye-tracking task (Sally-Anne-Task). In cases in which the number of passers and failers do not add up to the original number of subjects within each age group there were participants who did not provide an answer.

For the younger experimental group, children performed significantly below chance on the Smarties *other* question (25% correct,  $p < .05$ ). Performances did not differ from chance in the Smarties *self* task (42% correct,  $p = .54$ ) and in the prediction task in the last FB-trial (32% correct,  $p = .11$ ). For the older experimental group, the comparison between passers and failers in the Smarties *self* task yielded a significant difference: There were more passers than failers in this age group (87% correct,  $p < .001$ ). In contrast, performances on the other explicit tasks were at chance (both  $p$ 's  $> .05$ ). 5-year-olds showed significantly more correct answers in both Smarties tasks (*other*: 90% correct,  $p < .001$ ; *self*: 82% correct,  $p = .001$ ). Regarding the prediction question at the end of the eye-tracking task however, there was no difference between passers and failers for the control children, but for the adults (93% correct,  $p < .001$ ).

For all groups of children, there was no systematic performance difference between Smarties-*self* and Smarties-*other* task (McNemar test: all  $p$ 's  $> .05$ ). Younger 3-year-olds did not show a difference between each type of Smarties question and the prediction question at the end of the eye-tracking task either. There was a significant difference between the numbers of participants passing the Smarties *self*

**Table 6.2:** Performance on the different explicit FB tasks split by age group: Number (percentage) of participants who passed. *Note.* \* $p \leq .05$ , \*\*\* $p \leq .001$ .

Task	3;0–3;6 y total N = 25	3;7–4;0 y total N = 23	5-years old total N = 29	adults total N = 28
Smarties <i>other</i>	<b>6 (25%)*</b>	15 (65%)	<b>26 (90%)*</b>	—
Smarties <i>self</i>	10 (42%)	<b>20 (87%)*</b>	<b>23 (82%)*</b>	—
Sally-Anne (last FB trial)	8 (32%)	9 (39%)	10 (37%)	<b>25 (93%)*</b>

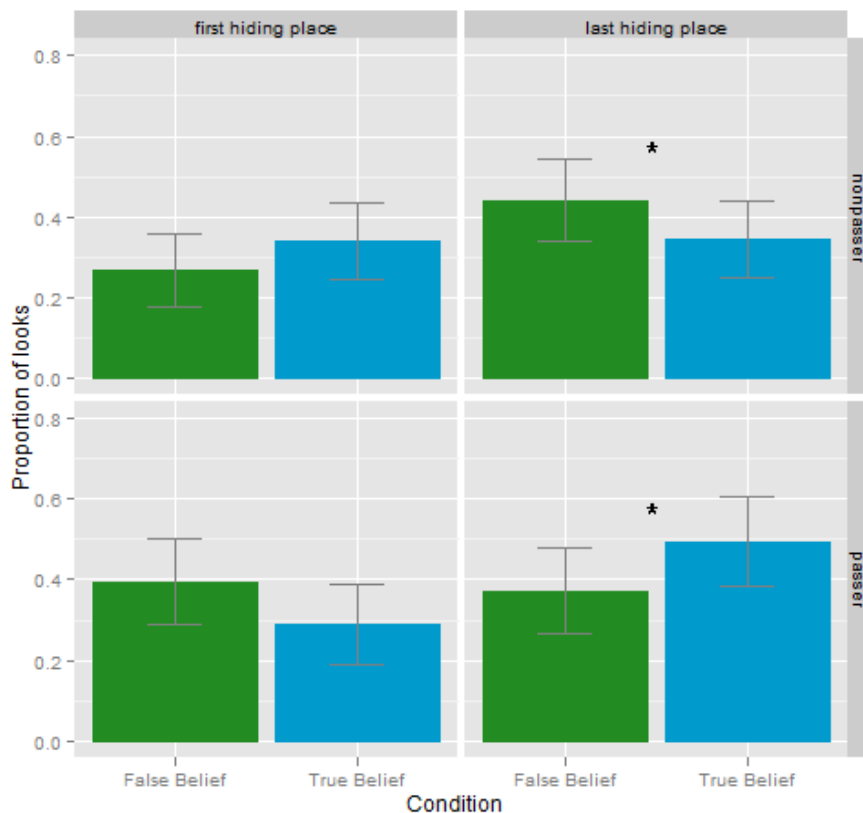
task and the prediction question for the older experimental group and 5-year-olds, i.e. children who succeeded only in one of both tasks were more likely to answer the Smarties *self* question correctly compared to performances on the prediction question (3;7–4;0-year-olds: 11 out of 20 Smarties *self*-passers failed the prediction task, whereas none of the nine prediction question-passers failed in the Smarties *self* task,  $p = .001$ ; 5-year-olds: 14 out of 22 Smarties *self*-passers failed the prediction task, but only 2 out of 10 prediction task-passers failed the Smarties *self* task,  $p < .01$ ). In addition, for the 5-year-olds the Smarties *other* task was significantly easier than the prediction task: 16 out of 25 Smarties *other* passers answered incorrectly in the prediction task, but only 1 out of 10 prediction question-passers failed the Smarties *other* task ( $p < .001$ ).

Analyses of the performance between age groups within each task yielded the following results: For the Smarties *self* question, the passer/failer ratio was significantly different between the younger and the older experimental group ( $\chi^2 = 10.43$ ,  $df = 1$ ,  $p < .01$ ) and between the younger experimental group and the 5-year-olds ( $\chi^2 = 9.13$ ,  $df = 1$ ,  $p < .01$ ). Performances of the older experimental group did not differ from the 5-year-olds ( $\chi^2 = 0.22$ ,  $df = 1$ ,  $p = .64$ ). Concerning the Smarties *other* task, children of the younger experimental group performed significantly worse than children of the older experimental group ( $\chi^2 = 7.68$ ,  $df = 1$ ,  $p < .01$ ) and than the 5-year-olds ( $\chi^2 = 22.95$ ,  $df = 1$ ,  $p < .001$ ). 5-year-old children outperformed the older 3-year-olds significantly ( $\chi^2 = 4.59$ ,  $df = 1$ ,  $p < .05$ ).

To summarize, 3-year-old children did not demonstrate FB understanding when predicting an upcoming action of another person: children younger than 3 years and 6 months performed below chance, whereas children aged between 3;6 and 4;0 years performed at chance. This is consistent with previous findings on explicit ToM task performance (Wellman et al., 2001). In addition, 3-year-olds performed better when attributing a former FB to themselves. The question at the end of the eye-tracking task did not elicit correct predictions even in 5-year-olds. This suggests that children were surprised by the question and thus answered rather randomly.

### 6.6.1.3 Relation between explicit and implicit ToM

In order to determine the relation between implicit and explicit ToM task performances in 3-year-olds, eye-tracking data were grouped according to explicit ToM task performance. An ANOVA with AOI (first- vs. last-hiding place) and condition (FB vs. TB) as within-subject factors was conducted for both Smarties-passers (N=21) and Smarties-nonpassers (N=26)<sup>5</sup>. Figure 6.9 illustrates the average fixation proportions at the first- and last-hiding place of each group and condition.



**Figure 6.9:** Average looking proportion towards first and last-hiding place in FB trials (green panels) and TB trials (blue panels); separate for Smarties task nonpassers (upper row) and Smarties task passers (lower row); error bars indicate standard errors for proportions. *Note.*  $*p \leq .05$ .

<sup>5</sup>Answers of the Smarties *other* question was taken as the core metric of explicit ToM performance because it is considered as the critical FB question as it taps the child's ability to attribute a FB to another person. In addition, approximately equal group sizes were achieved.

The interaction between *group* x *condition* reached significance for the last-hiding place ( $F(1, 43) = 7.77; p < .01$ ), and marginally for the first-hiding place ( $F(1, 43) = 3.44; p = .07$ ). For the passers, follow-up t-tests (one-tailed) between the FB condition and the TB condition revealed that they looked significantly longer to the last-hiding place in the TB condition compared to the FB condition (.49 vs .37,  $t(20) = 1.89; p < .05$ ). In contrast, nonpassers looked significantly longer towards the last-hiding place in the FB condition compared to the TB condition (.44 vs .34,  $t(23) = 2.05; p < .05$ ). The difference between conditions was not significant for the first-hiding place (passers: .29 vs .39,  $t(20) = 1.32; p = .20$ ; nonpassers: .27 vs .34,  $t(23) = 1.29; p = .21$ ). Overall, the looking behavior of the Smarties-passers was different from nonpassers (three-way interaction effect *group* x *condition* x *AOI*:  $F(1, 43) = 6.76, p = .01$ ): Whereas Smarties-passers displayed a looking pattern similar to that of 5-year-old children, nonpassers looked in FB-trials rather incorrectly. This indicates that implicit and explicit ToM task performances are associated (see discussion).

## 6.6.2 Language Measures

### 6.6.2.1 General Language Measures

Table 6.3 shows the average performance (with standard deviations in brackets) on TROG-D and AWST-R for all three groups of children. As for TROG-D, scores refer to the mean number of correct blocks (max. number of blocks: 21). For the AWST-R, the scores refer to the number of correct lexical items (max. 75). Expectedly, children's number of correct blocks (TROG-D) and number of correct labeled nouns and verbs (AWST-R) increased with age: Two-sample t-tests yielded significant differences in mean scores between the younger and the older experimental group (TROG-D:  $t(44) = -2.89, p < .01$ ; AWST-R:  $t(45) = -7.13, p < .001$ ), between the older 3-year-olds and the 5-year-olds (TROG-D:  $t(42) = -7.22, p < .001$ ; AWST-R:  $t(44) = -5.23, p < .001$ ) and between the younger 3-year-olds and the 5-year-olds (TROG-D:  $t(42) = -11.78, p < .001$ ; AWST-R:  $t(43) = -12.42, p < .001$ ).

To investigate whether both general language measures are correlated with each other, a partial correlation analysis, controlling for age, was conducted for 3-year-old children. There was a significant positive relation between TROG-D and AWST-R performances ( $r = 0.55, p < .001, df = 43$ ). That is, even when age (in months) was partialled out, general grammar comprehension increased as active vocabulary size increased.

**Table 6.3:** Performance on TROG and AWST-R split by age group [standard deviation in brackets].

	3;0–3;6 y	3;7–4;0 y	5-year-olds
TROG-D	5.74 [2.8]	8.43 [3.49]	15.48 [2.93]
AWST-R	29.79 [7.64]	46.26 [8.19]	58.43 [7.80]

### 6.6.2.2 Complementation Tasks

Table 6.4 presents the mean proportion (and standard deviations) of correct responses on the truth-value-judgment task, the sentence-picture matching task and the reporting-the-complement task as a function of age group. Firstly, performance on each task will be reported separately. To examine whether each age group performed significantly above chance on each task, one-sample Wilcoxon tests (two-tailed) were applied. Secondly, to compare performance between age groups Mann-Whitney U-Tests (two-tailed) were used. Partial correlation analyses were conducted within 3-year-olds to determine if age-unrelated relationships existed between each single task and general language measures. Therefore, age (in months) was included as control variable. Finally, correlations between the complementation tasks are reported, also adjusted for age effects.

**Truth-Value-Judgment Task** The mean overall performance, i.e. the average of correct acceptances of matching sentences *and* correct rejections of mismatching sentences, and the mean proportion of correct acceptances differed significantly from chance (.50) in all three age groups (OVERALL: 3;0–3;6 y:  $Z = 3.84, p < .001$ ; 3;7–4;0 y:  $Z = 4.16, p < .001$ ; 5-year-olds:  $Z = 5.16, p < .001$ ; ACCEPTANCES: 3;0–3;6 y:  $Z = 4.36, p < .001$ ; 3;7–4;0 y:  $Z = 4.12, p < .001$ ; 5-year-olds:  $Z = 5.18, p < .001$ ). However, children of the younger experimental group (3;0–3;6 years) rejected non-matching sentences in approximately half of the cases on average (.56) and thus performed at chance ( $Z = .725, p = .47$ ), whereas older 3-year-olds and 5-year-olds correctly rejected non-matching sentences significantly above chance (3;7–4;0 y: .85,  $Z = 3.76, p < .001$ ; 5-year-olds: .99,  $Z = 5.30, p < .001$ ). Since in TVJTs young children show a bias to give positive answers (Crain and Thornton, 1998), the correct no-responses are considered to be a more valid measure. Therefore, the following analyses on the TVJT refer to correct NO-answers only.

There were significant age effects across the three age groups tested: 3;0–3;6-year-old children performed significantly worse on correct rejections than 3;7–4;0-year-olds ( $U = 159, Z = 2.87, p < .01$ ) and 5-year-olds ( $U = 121, Z = 4.99, p < .001$ ). In addition, the 5-year-old children outperformed the older 3-year-olds significantly ( $U = 159, Z = 2.42, p < .05$ ).



Results of partial correlation analyses within 3-year-olds revealed a significant age-unrelated correlation between correct rejections and TROG-D performance ( $r = 0.433, p < .01, df = 43$ ) and between correct rejections and AWST-R ( $r = 0.358, p < .05, df = 43$ ).

**Sentence-Picture Matching Task** Given that there is a 33% chance to point to the correct picture in this task, children of the younger experimental group performed at chance level (37%,  $Z = 1.6, p = .10$ ) but the older 3-year-olds (58%) and the 5-year-olds (83%) performed significantly better than chance ( $Z = 3.5, p < .001, Z = 4.8, p < .001$ ). All age groups differed significantly from each other: the 5-year-old children performed significantly better than the younger 3-year-olds ( $U = 54, Z = 5.54, p < .001$ ) and the older 3-year-olds ( $U = 165, Z = 3.3, p = .001$ ). In addition, the older 3-year-olds outperformed the younger experimental group significantly ( $U = 157, Z = 2.9, p < .01$ ).

Comparisons between the two types of error (i.e. pointing to the unrelated picture vs. the embedded-clause-depicting picture) showed no differences between distractors within each age group: 3;0–3;6-year-olds selected the unrelated picture in 31% of the time and the embedded-clause picture in 30% of the time ( $Z = 0.09, p = .9$ ), 3;7–4;0-year-olds pointed more often to the picture that depicted the complement clause (23.9%) compared to the unrelated distractor (17.4%). However, this difference did not reach significance ( $Z = 0.88, p = .38$ ). This holds also for 5-year-olds (8.6% vs 6.9%,  $Z = 0.46, p = .64$ ).

A partial correlation analysis for the 3-year-olds revealed that this task significantly correlated, independent of age, with TROG-D ( $r = 0.34, p < .05, df = 46$ ), but not with AWST-R ( $r = 0.157, p = .30, df = 47$ ).

**Reporting-the-Complement Task** As presented in Table 6.4, the 3;0–3;6-year-old children performed in both the 2-representation-condition (false complements) and the 1-representation-condition (true complements) at chance level (false: .52,  $Z = 1.17, p = .24$ ; true: .56,  $Z = 0.76, p = .45$ ). Consequently, the overall performance did not differ from chance as well (.54,  $Z = 0.29, p = .77$ ). In addition, there was no condition effect for this age group ( $p = .42$ ). 3;7–4;0-year-olds showed a different response pattern: averaged across all items, they performed significantly above chance (.66,  $Z = 2.58, p = .01$ ). That is, however, not significantly different from the overall score of the younger age group ( $U = 193.5, Z = 1.65, p = .10$ ). There is a significant difference between conditions ( $Z = 2.14, p < .05$ ): While this age group gave significantly more correct responses in the 1-representation-condition compared to chance (.72,  $Z = 2.6, p < .01$ ), the mean score of the 2-representation-condition was at chance level (.59,  $Z = 0.68, p = .49$ ). 5-year-olds had high scores in both conditions, yielding above-chance performance in the 1-representation-condition

**Table 6.4:** Mean proportion and standard deviations of correct responses on the complement tasks: Truth-Value Judgment Task (TVJT), Sentence-Picture Matching Task (SPMT) and Reporting-the-Complement Task (RtCT), split by age group. *Note.* above-chance performance denoted by \*\* $p \leq .01$ , \*\*\* $p \leq .001$ .

	TVJT			SPMT		RtCT		
	correct acceptance	correct rejections	overall	false complements	true complements	overall		
<b>3;00-3;06 y</b> (N=25)	<b>.89***</b> (.20)	.56 (.40)	<b>.74***</b> (.20)	.37 (.16)	.52 (.22)	.56 (.27)	.54 (.21)	
<b>3;07-4;00 y</b> N=23	<b>.93***</b> (.21)	<b>.85***</b> (.29)	<b>.89***</b> (.17)	<b>.58***</b> (.28)	.59 (.28)	<b>.72**</b> (.24)	<b>.66**</b> (.22)	
<b>5-year-olds</b> N=29	<b>.96***</b> (.15)	<b>.99***</b> (.03)	<b>.98***</b> (.08)	<b>.83***</b> (.22)	<b>.87***</b> (.16)	<b>.89***</b> (.18)	<b>.88***</b> (.13)	

(.89,  $Z = 4.76, p < .001$ ), 2-representation-condition (.87,  $Z = 4.87, p < .001$ ) and consequently averaged across all 6 items (.88,  $Z = 4.73, p < .001$ ).

A comparison of performances between age groups revealed an expected effect: the 5-year-old children performed significantly better than the 3-year-old children in both conditions (3;0–3;6 yrs<sub>false</sub>  $U = 86, Z = 4.92, p < .001$ ; 3;0–3;6 yrs<sub>true</sub>  $U = 106, Z = 4.58, p < .001$ ; 3;7–4;0 yrs<sub>false</sub>  $U = 146, Z = 3.78, p < .001$ ; 3;7–4;0 yrs<sub>true</sub>  $U = 199.5, Z = 2.79, p < .01$ ). Interestingly, the experimental groups showed a marginally significant age effect for the 1-representation-condition ( $U = 186.5, Z = 1.91, p = .056$ ), but the younger and the older 3-year-olds did not differ in their scores in the 2-representation-condition.

Results of partial correlation analyses within 3-year-olds revealed a significant age-unrelated correlation between the overall performance and the TROG-D performance ( $r = 0.361, p < .05, df = 42$ ), but not between the overall performance and the AWST-R ( $r = 0.038, p = .81, df = 42$ ).

**Summary** In sum, children aged between 3;0 and 3;6 years did not master any of the complementation tasks<sup>6</sup>. This indicates that children at this age do not comprehend yet that a clause can be taken as an argument of a matrix verb, and thus fail to process the dependency of the subordinate clause. Older 3-year-olds showed complement clause understanding in the truth-value-judgment task, the picture-sentence matching task and additionally in the 1-representation-condition of the reporting-the-complement task. However, they seem to be limited to true complements since they still performed at chance when reporting false statements in the reporting-the-complement task. As expected, the 5-year-old control children successfully passed each task including the false-complement subtest and hence showed a full-fledged complement sentence comprehension. With respect to the results of the correlation analyses, both the sentence-picture matching task and the reporting-the-complement task correlated with the general syntax score (TROG-D), but not with productive vocabulary (AWST-R) whereas the truth-value-judgment task correlated with both. This suggests that while the former tasks tapped syntactic knowledge without overlapping with vocabulary size, the later task tapped rather a broader language ability.

**Correlations between complementation tasks** Since all three complementation tasks were designed to measure the same construct, namely the comprehension of complement sentences, high correlations between them are expected. In order to avoid spurious correlations due to ceiling performances in the 5-year-olds only the data from the 3-year-olds were included into the correlational analysis. All

---

<sup>6</sup>The good performance in accepting matching sentences in the truth-value-judgment task is assumedly due to a yes-response bias.

screening tasks were correlated with age in months: truth-value-judgment task:  $r = 0.48, p < .001$ ; sentence-picture matching task:  $r = 0.39, p < .01$ ; reporting-the-complement task:  $r = 0.36, p < .01$ . Table 6.5 shows the partial-correlations among the screening tasks. Controlling for age, significant correlations (one-tailed) were found between the sentence-picture matching task and all three measures of the reporting-the-complement task (overall:  $r = 0.411, p < .01$ ; false complements:  $r = 0.318, p < .05$ ; true complements:  $r = 0.355, p < .01$ ) as well as between the two subtests of the reporting-the-complement task ( $r = 0.342, p < .05$ ). As a matter of fact, both subtest scores of the reporting-the-complement task (false and true complements) are highly significantly correlated with the overall RtCT score. Interestingly, the truth-value-judgment task is not correlated with the other two screening tests when the effect of age is controlled for (all  $p$ 's  $> .57$ ).

**Table 6.5:** Partial correlations between screening tests controlling for age, (3-year-old children only). *Note.*  $*p \leq .05$ ,  $**p \leq .01$ ,  $***p \leq .001$ .

	TVJT	SPMT	RtCT		
			overall	false complements	true complements
TVJT correct rejections	1	-	-	-	-
SPMT	0.088	1	-	-	-
RtCT overall	0.059	<b>0.411**</b>	1	-	-
RtCT false	0.02	<b>0.318*</b>	0.812***	1	-
RtCT true	0.76	<b>0.355**</b>	0.826***	<b>0.342*</b>	1

### 6.6.3 Relations between FB tasks and language tests

#### 6.6.3.1 Explicit ToM and language measures

**Correlational analyses** The following analyses included only the 3-year-olds. Bivariate correlations between Smarties *self*, Smarties *other*, AWST-R, TROG-D, sentential complementation screening tests and age in months are reported in Table 6.6.

It was found that both Smarties *self* and Smarties *other* significantly correlated with age (SELF:  $r = 0.38, p < .01$ ; OTHER:  $r = 0.29, p = .05$ ), AWST-R (SELF:  $r = 0.49, p < .001$ ; OTHER:  $r = 0.53, p < .001$ ), TROG-D (SELF:  $r = 0.34, p < .05$ ;

**Table 6.6:** Bivariate correlations among explicit ToM measures, age and vocabulary size (AWST-R), general syntax (TROG-D) and sentential complement tasks (3-year-olds only). *Note.* \* $p \leq .05$ , \*\* $p \leq .01$ , \*\*\* $p \leq .001$ .

	Smarties		Age	AWST-R	TROG-D	TVJT		RtCT	
	<i>self</i>	<i>other</i>				<i>corr. rejections</i>	SPMT	<i>overall</i>	<i>false true</i>
Smarties <i>self</i>	1	-	-	-	-	-	-	-	-
Smarties <i>other</i>	<b>.46***</b>	1	-	-	-	-	-	-	-
Age	<b>.38**</b>	<b>.29*</b>	1	-	-	-	-	-	-
AWST-R	<b>.49**</b>	<b>.53***</b>	<b>.64***</b>	1	-	-	-	-	-
TROG-D	<b>.34*</b>	<b>.30*</b>	<b>.45**</b>	<b>.66***</b>	1	-	-	-	-
TVJT corr. rej.	<b>.37*</b>	<b>.37*</b>	<b>.48***</b>	<b>.54***</b>	<b>.55***</b>	1	-	-	-
SPMT	.22	<b>.32*</b>	<b>.39**</b>	<b>.41**</b>	<b>.48***</b>	<b>.30*</b>	1	-	-
RtCT overall	.06	-.02	<b>.36*</b>	.27	<b>.47***</b>	.24	<b>.50***</b>	1	-
false c.	.11	-.05	.28	.09	<b>.33*</b>	.16	<b>.39**</b>	<b>.83***</b>	1
true c.	-.01	.02	<b>.32*</b>	<b>.37*</b>	<b>.46***</b>	.23	<b>.44***</b>	<b>.85***</b>	<b>.40**</b>

OTHER:  $r = 0.3, p < .05$ ) and the truth-value-judgment task (SELF:  $r = 0.37, p = .01$ ; OTHER:  $r = 0.37, p = .01$ ). In addition, there was a significant correlation between Smarties *other* and sentence-picture matching task ( $r = 0.32, p < .05$ ).

**Logistic regression** A hierarchical logistic regression analysis was conducted in order to determine whether sentential complement understanding predict explicit ToM task performance after controlling for age. Note that the number of predictors is limited because of the small sample size in the present study (see Backhaus et al., 2008; Peduzzi et al., 1996). Therefore, a composite score for sentential complement mastery (COMPscore) was calculated by averaging across scores of the sentence-picture matching task and truth-value-judgment task (N=48, mean: 58.54, std: 25.39, min: 12.5, max: 100). The reporting-the-complement task was not included in this composite score because of its weak correlation with explicit ToM measures.

As can be seen in Table 6.7, on the first step *age* was entered as independent variable and Smarties *self* (passer/nonpasser) (model 1, see A) and Smarties *other* (model 2, see B) as the dependent variable. The COMPscore was then entered on the second step. As expected, age was a significant predictor of Smarties *self* ( $Wald(1) = 5.8, p < .05$ ;  $oddsratio[OR] = 1.324$ ) and Smarties *other* ( $Wald(1) = 3.74, p = .05$ ;  $OR = 1.211$ ) in the first step of both models. The model (step 1) including age correctly predicted explicit ToM performance status in 68.1% of the cases (Smarties *self*) and 66% cases (Smarties *other*)<sup>7</sup>. The contribution of sentential complement comprehension in the second step differed between the two models. As for the Smarties *self* task, the COMPscore made no significant contribution to prediction of (non)passing status ( $Wald(1) = 1.90, p = .17$ ) and also age failed to be a significant predictor when simultaneously entering the COMPscore ( $Wald(1) = 1.89, p = .17$ ). In terms of model evaluation, the second step did not improve the model significantly ( $\chi^2(1) = 1.96, p = 0.16$ ). Concerning the Smarties *other* task, the COMPscore was found to be a significant predictor ( $Wald(1) = 4.86, p < .05, OR = 1.04$ ), but again age failed to reach significance in the second step ( $Wald(1) = 0.26, p = .61$ ). This pattern is indicative for a mediator effect, that is the relation between age and Smarties *other* (as shown in the first step) is fully explained by the relation between complement mastery and the explicit ToM measure. Furthermore, the second step improved the model significantly ( $\chi^2(1) = 5.67, p < 0.05$ ).

---

<sup>7</sup>The cutoff was set 0.5 in the classification table.

**Table 6.7:** Results of hierarchical logistic regression analyses for sentential complement understanding (composite score) predicting Smarties task performance (self and other), controlling for age (3-year-old children only).

<b>A) Smarties <i>self</i></b>						
	$\beta$	$SE \beta$	Wald	$df$	$p$	odds ratio
step 1						
constant	<b>-11.38</b>	4.89	5.31	1	.021	.000
age	<b>0.28</b>	0.12	5.80	1	.016	1.324
step 2						
constant	-8.48	5.31	2.55	1	.110	.000
age	0.18	0.13	1.89	1	.168	1.203
COMPscore	0.02	0.02	1.89	1	.168	1.023
<b>B) Smarties <i>other</i></b>						
	$\beta$	$SE \beta$	Wald	$df$	$p$	odds ratio
step 1						
constant	<b>-8.39</b>	4.25	3.89	1	.048	0.000
age	<b>0.19</b>	0.10	3.74	1	.053	1.211
step 2						
constant	-5.08	4.49	1.28	1	.258	0.006
age	0.06	0.11	0.26	1	.607	1.060
COMPscore	<b>0.04</b>	0.02	4.86	1	.027	1.040

**Pass/fail contingencies** To test the linguistic determinism hypothesis as de Villiers proposed it, a closer look at the scores of the reporting-the-complement task including false complements is essential. The weak correlation between the reporting-the-complement task and the explicit FB task might be due to children who performed well on the complementation task but still failed the Smarties task. These cases, however, would not contradict the linguistic determinism hypothesis; only children who pass the ToM task without comprehending sentential complements would challenge the syntax-first hypothesis.

Therefore, a contingency analysis was conducted to test for the direction of the effect between the reporting-the-complement task (false complements) and the Smarties *other* task. The criterion for passing the reporting-the-complement task was set at 2 or 3 correct items out of 3 items. This is a rather lenient criterion but it reflects the fact that children's performance is not compared to chance (50%) but with what they would assumedly do if they did not comprehend sentential complementation at all, that is reporting the proposition of the simple sentence and thus responding incorrectly. In addition, even for 5-year-olds who displayed a highly significant above-chance performance in all complementation tasks only 18 children (62%) would be counted as passers if the "passing" criterion was set at 3 correct out of 3 possible.

Table 6.8 presents the number of 3-year-olds who simultaneously passed, simultaneously failed, passed the Smarties *other* but not the reporting-the-complement task, and passed the reporting-the-complement task but not the Smarties task.

**Table 6.8:** Contingency table showing numbers of children passing and failing the Smarties *other* task and the reporting-the-complement task (false complements only) (3-year-old children only).

		RtCT <i>false</i>	
		Pass	Fail
Smarties <i>other</i>	Pass	14	6
	Fail	15	10

The crucial question is whether the number of children who fell above the diagonal (i.e. Smarties passers/ language failers) or below the diagonal (i.e. Smarties failers/ language passers) is equally distributed or not. The number of children who passed the Smarties but not the language task (N=6) is significantly lower than the number of children who passed the language but not the Smarties task (N=15) (binomial test,  $p < .05$ , one-tailed). In addition, considerations of the other language measures of the six "exceptional" children revealed that four of them correctly extracted true complements and thus passed one subtest of the reporting-the-complement task<sup>8</sup>. That indicates that these children comprehended at least the syntactic property of sentential complements.

To examine the empirical link between explicit ToM and children's performance on extracting true complements, the same contingency analysis was employed for the 1-representation condition of the reporting-the-complement task (see Table 6.9). Again, there were significantly more children who failed the Smarties *other* task but showed linguistic evidence of true complement comprehension (N=19) than children who showed the reverse performance pattern (N=4) ( $p < .01$ , one-tailed). Two out of these four "exceptional" children succeeded in the 2-representation subtest.

<sup>8</sup>Criterion for passing was again 2 or 3 correct items out of 3.



**Table 6.9:** Contingency table showing numbers of children passing and failing the Smarties *other* task and the reporting-the-complement task (true complements only) (3-year-old children only).

		RtCT <i>true</i>	
		Pass	Fail
Smarties <i>other</i>	Pass	16	4
	Fail	19	6

The finding that the same direction of effect was obtained also for true complement comprehension suggests that comprehending falsity is not essential for passing the Smarties *other* task. However, there was an overlap of children who passed both the true and the false complement condition (N=25) which is also reflected in the correlation coefficient of both subtests. To test whether syntactic knowledge of complement sentences alone is sufficient for passing the ToM task, the explicit ToM performance of children who passed *only* the true complement subtest (N=10) was compared to that of the children who solved both subtests (N=25). A chi-square test yielded no significant difference: The passer/failer ratio was equal for children who passed both subtests (12 passers/ 13 failers) to those children who showed only evidence of true complement comprehension (4 passers/ 6 failers) ( $\chi^2 = 0.18, df = 1, p = .62$ ). That is, there is no evidence that comprehending false complement clauses marks an additional step that is essential for the relational link between FB understanding and the comprehension of sentential complements.

### 6.6.3.2 Implicit ToM and language measures

Since the eye-tracking data did not reveal an early implicit FB understanding in 3-year-olds, the implications of the following analyses are limited. However, there were interindividual differences in looking proportions (as indicated by the dispersion) which may be explained by interindividual differences in language comprehension.

Partial correlation analyses, controlling for age, were carried out to detect potential empirical connections between the 3-year-olds' looking behavior in the FB condition and their language tasks performance. Two metrics were used to quantify correct looking: (1) averaged looking time difference (DIFFscore) within the FB condition (i.e. looking at first-hiding AOI – looking at last-hiding AOI), and (2) mean proportion of correct anticipations (AOIcor) calculated by dividing the looking time towards the first-hiding AOI by the sum of looking time to both AOIs. Thus, AOIcor can range between 0% and 100%, being 50% if there were equal anticipations towards both AOIs, and 100% if there were solely correct

anticipations. Inherently, both variables are highly related (Pearson correlation,  $r = 0.91, p < .001, df = 46$ ). However, no correlations were found between measures reflecting implicit knowledge and sentential complement screening tests (see Table 6.10<sup>9</sup>). Interestingly, a significant negative partial correlation was found between TROG-D and DIFFscore ( $r = -0.409, p < .01$ ) and AOIcor ( $r = -0.503, p = .001$ ). However, negative partial correlations between AWST-R and each implicit ToM metric did not reach significance ( $p > .06$ ).

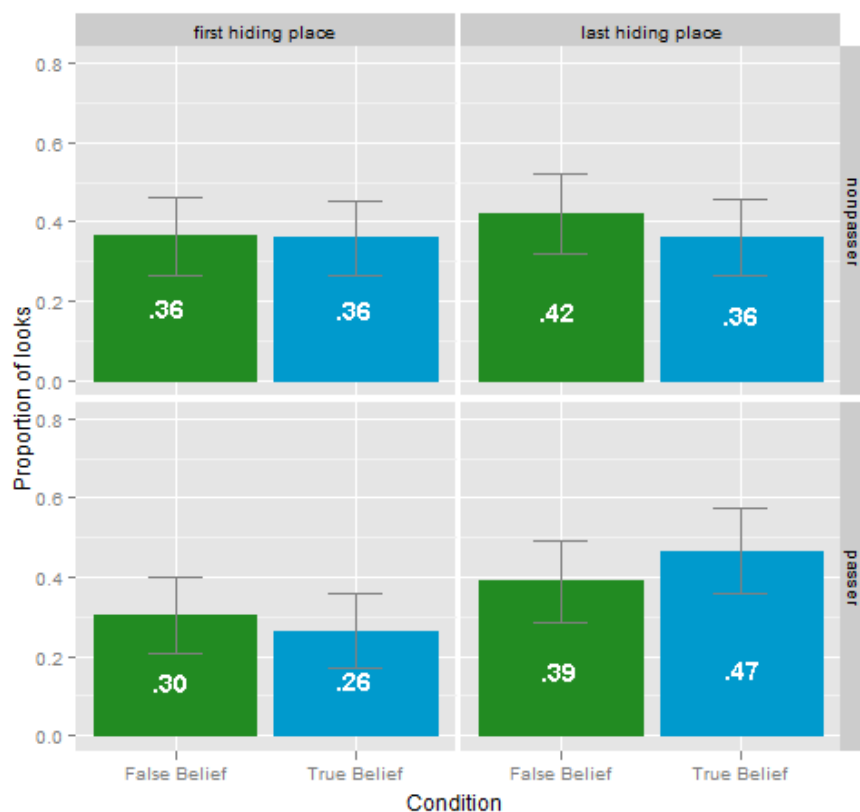
**Table 6.10:** Partial correlations between implicit ToM metrics and language tests controlling for age (3-year-old children only). *Note.*  $*p \leq .05$ ,  $**p \leq .01$ ,  $***p \leq .001$ .

	DIFFscore	AOIcor	TVJT	SPMT	overall	AWST-R	TROG-D
DIFFscore	1	.905***	-.104	-.011	-.049	-.224	<b>-.409**</b>
AOIcor	.905***	1	-.162	-.108	-.096	-.282	<b>-.503***</b>

As a second approach, eye-tracking data were grouped according to the language tasks performance of the 3-year-olds. Therefore, a composite score was calculated across the three complement tasks (sentence-picture matching task, truth-value-judgment task and reporting-the-complement task, overall) (N=48, mean=59.90, median=61.11, std=21.35, min=19.44, max=100). A mediansplit method was used to classify participants as complement-nonpassers (N=25, mean=41.66, std=13.54, min=19.44, max=61.11) and complement-passers (N=23, mean=77.66, std=8.27, min=63.89, max=100).

An analysis of variance (ANOVA) with AOIs (first- vs. last-hiding place) and condition (FB vs. TB) as within-subject factors and group (complement-passers vs. complement-nonpassers) as between-subject factor was conducted. Figure 6.10 illustrates the average fixation proportions at the first- and last-hiding place of each group and condition. Overall, the looking behavior of the complement-passers was not significantly different from complement-nonpassers (three-way interaction effect *group x condition x AOI*:  $F(1, 44) = 1.22, p = .27$ ). In a separate 2 x 2 (group x condition) ANOVA within the last-hiding place, the interaction between group and condition did not reach significance ( $F(1, 44) = 2.89, p = .10$ ). This analysis confirms that complement mastery and anticipatory looking in the predictive looking task were not associated (see discussion).

<sup>9</sup>Note that correlations among language tests are not given here as they are shown in 6.5.



**Figure 6.10:** Average looking proportion towards first and last-hiding place in FB trials (green panels) and TB trials (blue panels); separate for complement tasks nonpassers (upper row) and complement tasks passers (lower row); error bars indicate standard errors for proportions.

## 6.7 Discussion

The discussion on *Study 1* is divided into three parts corresponding to the research questions that were addressed. First, children's performance on a nonverbal implicit version of a FB task was examined. This task differed from previous used indirect FB tasks as it was completely nonverbal *and* it was as complex as the traditional change-of-location task. Second, three different language tasks were conducted to tap sentential complement comprehension in German. Besides children's general mastery of the grammar of complementation, it was tested whether understanding false complements is more complex than understanding true complements. The third and central question concerns the relation between sentential complement understanding and both explicit and implicit FB understanding. Whereas a developmental link between complement sentence comprehension and understanding beliefs has been found for

explicit ToM in previous research (e.g., de Villiers and Pyers, 1997, 2002; Low, 2010), the relation between implicit ToM (measured by eye gaze) and complement mastery is relatively unexplored.

**Validation of the nonverbal implicit FB task** Contrary to the hypothesis and to previous research, data on children's anticipatory eye gaze did not reveal any early FB sensitivity: Irrespective of the protagonist's belief, 3-year-old participants expected the story's character to show up at that location where the object really was. 5-year-old children and adults, however, displayed the expected gaze pattern and thus FB understanding, that is more anticipatory looks towards the place that the protagonist believed to find the object, irrespective of the real state of affairs. Therefore, the possibility can be ruled out that the task was simply an invalid means of measurement; rather, the material was appropriate in order to elicit looking behavior that reflects a spontaneous attribution of a mental state.

In a second approach, the relation between the implicit and the explicit FB measure was investigated to validate the current eye-tracking task. 3-year-old children's performance on the implicit FB task was associated with their explicit ToM task performance: whereas children who passed the explicit ToM task tended to display similar eye movement patterns as 5-year-old children did, children who failed the Smarties task looked rather incorrectly. This was not simply an age effect because separate analyses for younger and older 3-year-olds revealed that both age groups did not differ in their looking behavior on average.

The fact that 3-year-olds showed no earlier implicit FB sensitivity but instead explicit and implicit measure were associated casts doubts on the general assumption that implicit knowledge precedes explicit knowledge (Clements and Perner, 1994; Karmiloff-Smith, 1992). However, the current task has several properties that differ from previous implicit FB tasks and which may had an effect of children's visual orientation. First, the task was completely nonverbal; neither was verbal input given nor was a prompt question posed. Thus, the current task is most comparable with a task designed by Yi (2009) who also measured anticipatory eye movements and found no evidence for implicit FB understanding in 3- and even in 4-year-olds. In the original task conducted by Clements and Perner (1994), the FB story was narrated while it was enacted by the experimenter (this holds also for Clements et al., 2000; Low, 2010; Ruffman et al., 2001a). For example, the goal of the protagonist was stated explicitly before he started to disappear behind an occluder ("I'll go and get the cheese.").

In addition, there were control questions to check whether the child remembered the critical aspects of the story correctly (e.g, where the actor put the object first, where the object is now and whether the actor saw the displacement). Stories were repeated in case the control questions were answered incorrectly. Third, a prompt

was given to elicit anticipatory looks (e.g., *I wonder where he's going to look?*) which also highlights that it is not clear where the actor is going to search. Given that the current task did not include these three features, one possible explanation for the incorrect looking behavior of 3-year-olds is that children take advantage of verbal input to follow the story and/or the prompt question prior to the critical anticipation period helps them to focus their attention on the protagonist's intention. In addition, it is also possible that some participants in the present study did not understand the critical aspects of the story as no control questions were asked. Notably, in the current study 5-year-olds' expectations were not reflected in their first looks but only in their looking proportion data within the anticipation window of 1600ms. This supports the assumption that the current task was harder compared to the original one. The possibility that simply the video presentation in contrast to real objects and locations may had an effect on children's performance can be ruled out. Low (2010) presented the predictive looking task as both, a fixed video and an acting-out version and found no performance difference between both formats.

The consideration of the experimental setting in a second type of anticipatory FB task (Southgate et al., 2007) offers an alternative explanation. As in the current study, the stimuli were also presented nonverbally and as videos in the Southgate et al.'s (2007) study, but crucially the target object was removed from the scene before the protagonist started to reach for it. In the current task, the object was still present in one of two relevant locations at the time the participants anticipated the protagonist's reappearance. This enhances the likelihood of looking towards the object's actual place which is the incorrect location in the FB scenario. This *pull towards reality* effect is known for explicit FB tasks (cf. Wellman et al., 2001) and presumably plays a role in the current task, too. However, it was aimed to maintain the complexity of the standard FB task. It has been discussed controversially in the literature whether the presence of the target object is essential or not. Perner et al. (2007) argue that Southgate et al.'s tasks lacks *referential ambiguity*, that is there is no competition between the reality information and the actor's false representation of reality. Thus, children's correct looking behavior might result from applying a simple association strategy (e.g., looking-where-she-last-saw-it) rather than attributing a FB to the protagonist.

In sum, the findings suggests that the new implicit nonverbal FB task was suitable for testing FB understanding in general. The fact that implicit ToM was not found in 3-year-olds is assumedly due to the properties of the task. The task's complete nonverbal nature while maintaining the complexity of a traditional FB task possibly made the task more demanding compared to previously used predictive looking tasks. It is plausible that task-specific variables are influential and findings on implicit ToM therefore are constrained to a certain set of experimental settings as proposed also by Yi (2009).

**Comprehension of sentential complements** Findings on children's comprehension of complement sentences provide evidence that children aged between 3;0 and 3;6 years do not understand yet that a clause can serve as an argument of a complement-taking verb since they performed at chance level in each of the three tasks. Older 3-year-olds (mean age 3;9 years), however, displayed a target-like interpretation of complex sentences in all tasks. Yet, they did not perform at ceiling since 5-year-olds still outperformed them. Significant age-unrelated correlations with the receptive grammar task (TROG-D) confirmed that each task assessed syntactic knowledge. Only the truth-value-judgment task additionally correlated with vocabulary size (AWST-R) suggesting that this task tapped a broader language ability.

The second major finding was that extracting true complements was significantly easier than extracting false complements for older 3-year-old children. Previous research on the mastery of sentential complementation has mainly focused on children's ability to extract false complements. For example, the widely used memory-for-complement task involves children's understanding that the proposition of a complement clause need not to be true (e.g. used in de Villiers and Pyers, 2002; Hale and Tager-Flusberg, 2003; Lohmann and Tomasello, 2003; Low, 2010). To disentangle the syntactic and semantic properties of sentential complements children were asked to report both true and false complements in the reporting-the-complement task, a modified version of the memory-for-complement task. Whereas both older 3-year-old and 5-year-old participants significantly performed above chance when reporting true complements, only 5-year-olds mastered to report the content of a protagonist's false utterance. This is despite the fact that the true complement items were harder in terms of more ambiguous than items of the false complement condition because the embedded proposition was a paraphrase of the proposition of the simple sentence. That is, reporting the simple sentence instead of the complement clause in the true complement condition does not necessarily imply that the child was not able to process the complex sentence; rather, it is also possible that the participant simply rephrased the embedded clause. However, in the false complement condition children's incorrect answer points more strongly to an incompetence of complement comprehension. The reason for the above-chance performance of older 3-year-olds in the sentence-picture matching task and the truth-value-judgment task is presumably that both tasks required the child to understand the dependence of the embedded clause from the matrix clause without assessing the child's ability to understand falsity. In sum, results strongly suggest that children start to comprehend the syntactic complexity of complement sentences during their fourth year of age, but this knowledge seems to be restricted to sentences where the embedded argument has the same truth value as the main clause.

The finding on false complement performance is in line with de Villiers and Pyers' (2002) result who found that only about 58% of children aged 3;9 on average

passed the original memory-for-complement task. In contrast, Thornton and Crain (1994, Exp. 3) demonstrated that already children between 3;0 and 4;1 correctly interpreted long distance wh-questions in which the child is required to extract what the protagonist mistakenly said (e.g., What did he say's under there? while the puppet guessed something wrong before). The authors thus concluded that children have the syntactical knowledge by the age of 3;0. A possible explanation for the opposite results could lie in the different experimental setups. While de Villiers and Pyers (2002) and the current study presented a story containing both the complement sentence expressing the protagonist's false representation of reality and the simple sentence expressing the reality, children in Thornton and Crain's experiment experienced the real state of affairs by manipulating it actively. Possibly, the more interactive paradigm without providing the reality information linguistically made the task less demanding. However, the setup in Thornton and Crain's experiment poses also the question what the child needs to process to respond correctly. Since the reality information was given visually and was thus shared by the experimenter and the child the only new information was what the puppet said. The child does not need necessarily process the complete complex question to respond correctly but only the main clause (*What did he say?*) because pragmatically it is felicitous to report the new information (see also Dąbrowska et al., 2009, for a similar argumentation). An examination of this assumption could involve a closer look on children's responses: if children simply processed the main clause instead of the complex question their answers should be more prone to entail the whole complement clause (e.g., *The object is under there.*). However, if they did comprehend the complement structure a minimal response (i.e. the object of the complement clause) is appropriate. Unfortunately, Thornton and Crain (1994) did not differentiate between these response types.

In general, results of the present work support the hypothesis of a stepwise acquisition of sentential complementation (de Villiers, 2005a). However, there are differences concerning the question what exactly it is that enables children to interpret false complements correctly. According to de Villiers (1999, 2005a), it concerns syntactic knowledge as children - in a final step - acquire the so-called Point-of-View (PoV) feature in the complementizer phrase. De Villiers proposes that in case the child's grammar lacks this feature, "the initial marking is for an obligatorily true proposition under CP, else the sentence is judged as "false" with respect to the world" (de Villiers, 1999, p.131). As language acquisition proceeds, children assumedly understand that the truth of the complement clause is relative to the subject of the matrix clause and not relative to the world. However, only a certain set of complement-taking verbs, namely nonfactive mental state and communication verbs, allow the representation of false propositions. This points towards an involvement of semantic knowledge. Indeed, Hale and Tager-Flusberg (2003) and Tager-Flusberg and Joseph (2005) refer to this unique property of

sentential complements as a semantic feature.

### **Relation between ToM tasks and sentential complement comprehension**

The primary aim of the study was to investigate the developmental relation between sentential complement comprehension and ToM, especially whether the linguistic determinism hypothesis also holds for implicit ToM. It was found that complement mastery in two out of three tasks coincides in time with successful explicit ToM task performance but not with children's anticipatory looking. This result suggests the existence of two separate representational systems as it demonstrates that explicit but not implicit FB understanding is related with children's understanding of complex structures. It is important to note that because implicit ToM could not be found in 3-year-olds the implications of the correlational analyses are restricted. Yet, associations between children's explicit FB task responses and their anticipatory looking behavior have been identified. Hence, at least theoretically, the possibility remains that interindividual differences in the predictive looking task are associated with different levels of sentential complement understanding. Strikingly, a negative correlation between general syntactic knowledge (TROG-D) and gaze has been found. However, the TROG-D covers a broad range of syntactic structures (relative clauses, prepositions, etc.). Therefore, it is not evident which aspect exactly caused the negative correlation.

As for explicit ToM, the present study replicated former findings on the developmental link between sentential complement understanding and traditional FB tasks (de Villiers and Pyers, 1997, 2002; Low, 2010). More interestingly, current results expand on previous work by revealing that the relationship also holds for tasks that assessed children's ability to process complement sentences without involving misrepresentation. Significant correlations were found between the Smarties task and both the sentence-picture matching task and truth-value-judgment task. Crucially, children were not required to evaluate the embedded proposition against reality in both tasks. Instead, the comprehension of the syntactic complexity of a complement sentence was tapped. A hierarchical logistic regression revealed that this correlation was not simply an effect of age (i.e., because both task develop concurrently) as the composite score of both language tasks explained variance after controlling for age. Importantly, the relationship between age and explicit ToM was no longer statistically different from zero after the complement score was entered into the model. This is indicative for a mediation effect, that is, the observed relationship between age and explicit ToM is explained completely by the relation between language and ToM (see Preacher and Hayes, 2004, for implementation and assumptions for testing mediation effects).

However, the analysis did not include general language measures as control



variables like it was done in previous studies (de Villiers and Pyers, 2002; Low, 2010; Tager-Flusberg and Joseph, 2005). Thus, it can not be excluded that general language ability and not specifically the knowledge of sentential complementation is responsible for the effect. Like all maximum likelihood procedures, a logistic regression requires a large sample sizes or the number of predictors has to be adjusted for smaller samples, respectively. To provide clearer evidence, one has to either increase the sample size or collect multiple measures of explicit FB understanding so that a linear regression analysis can be conducted.

The third sentential complementation task (reporting-the-complement task) did not correlate with the Smarties task, yet contingency analyses revealed that 3-year-olds were significantly less likely to answer the explicit FB task correctly while failing the language task than, reversely, passing the language task while failing the Smarties task. Hence, the linguistic determinism hypothesis is again supported by the results. Nevertheless, there were two 3-year-olds who failed both subtests of the reporting-the-complement task but predicted the upcoming behavior of a person who is holding a FB correctly. According to the linguistic determinism hypothesis, this is not possible. However, research on test-retest reliability of FB tasks suggests that young children's performance might be not stable (Mayes et al., 1996, but see Hughes et al., 2000 for different results). In other words, it cannot be ruled out that some children who possessed a ToM answered incorrectly and vice versa. Conducting additional explicit ToM tasks or retesting the children on the Smarties task, respectively, would minimize the risk of misclassification.



# Chapter 7

## Study 2: Implicit Understanding of Complement Sentences

### 7.1 Introduction/ Aims

In previous studies on the developmental link between FB understanding and comprehension of sentential complementation the tasks assessing children's syntactic knowledge like the memory-for-complement task require a conscious, verbal reaction from participants. As has been confirmed in *Study 1* of the current work, a full perceptive acquisition of this specific structure has been observed not before age 3;6. To get closer to the question if an implicit ToM found in children considerably younger than 3;6 develops independently from linguistic representation, *Study 2* takes a methodological approach. Two experiments were designed to investigate children's comprehension of sentential complements in a more implicit fashion by using the same means of measurement as used in experimental tasks tapping implicit FB understanding: anticipatory eye movements. The main aim thus was to find out whether this passive and task-free methodology reveals any precocious understanding of complement sentences.

In Experiment 1, 3-year-old children were presented with complex sentences expressing a protagonist's FB and a simple sentence mentioning the reality. This is analog to the memory-for-complement task involving mental state verbs and false complements. But instead of asking children to report the proposition of the embedded clause their expectations about the protagonist's future behavior were measured. The basis assumption is that children who are proficient in processing the complex sentence are guided by the linguistic representation when anticipating the protagonist's upcoming mistaken action. Besides the new method, this is interest-

ing also from a theoretical point of view: Given that 3-year-olds show already FB understanding in their visual orientation in tasks in which the FB story is enacted (i.e. the unseen displacement is shown) the experimental setup sheds light on the question of whether language comprehension feeds into the system that underlies implicit ToM. In other words, are children also affected in their anticipations when the information about a person's FB is provided solely verbally?

To tease apart effects of syntactic complexity from cognitive skills, Experiment 2 tested children's understanding of sentential complements without entailing contradicting propositions. That is, as in Experiment 1, children heard complement sentences expressing the content of a protagonist's belief but no information about whether this belief is true or not was given. To ensure that the processing of the matrix clause is crucial for the interpretation of the sentence and the looking behavior of the children is not solely directed by the lexical information, Experiment 2 focused on the scope of negation over a complement clause (see Methods).

Comparisons between both experiments allow to determine the impact of additional cognitive demands. In case performances differ between Experiment 1 and 2, this would indicate that the conflicting reality information in Experiment 1 masked an understanding of the syntactical complexity of complement sentences.

## 7.2 Experiment 1: Expressing a False Belief

### 7.2.1 Participants

Forty-one German-acquiring children aged between 3 and 4 years participated in the study (18 female). The group was split into a younger and an older experimental age group: 20 younger 3-year-olds (mean age: 3 years 4 months (3;4); range: 2;11 - 3;6, 10 female) and 21 older 3-years-olds (mean age: 3;10; range: 3;7 - 4;1, 8 female). In order to validate the new method, 16 5-year-old children (mean age: 5;4; range: 4;s 11 - 5;9; 6 female) and 19 adults (mean age: 27 years, age range: 21 - 42 yrs, 15 female) were tested. One child of the younger experimental group was excluded from the data analysis because of data loss during the experiment. All participants were recruited in the Berlin/ Potsdam area. According to a questionnaire which was filled in by the parents all children were developing normally and were growing up in monolingual homes. To check for red/green color blindness children were asked at the beginning to identify a red-colored and green-colored elephant on a painting. This was done because the colors were used in the eye-tacking task. All of the participants were able to discriminate the colors.

### 7.2.2 Material

The experiment contained eight nonverbal familiarization trials at the beginning and four experimental trials in total. In the familiarization trials, each participant was presented with the visual scene (see Figure 7.1/1) and with the critical event of the movie, i.e. the fact that the protagonist of the story (the mouse) starts to move away from the window and shows up again in that mouse hole that is next to the box that the mouse believes to find an object in. This object had been hidden by a cat in one of the two boxes. The hiding procedure had either been observed by the mouse (TB context) or not (FB context). The familiarization was completely nonverbal, i.e. there were no comments about the mental state of the protagonist. The purpose of the familiarization phase was to ensure that the participants knew that the mouse has access to both mouse holes and always wants to obtain the hidden object after it disappeared. Whether the mouse finds the object or not was not shown in the movies.

At the beginning of each of the four experimental trials three context sentences were presented that introduced the mouse and the hidden object as exemplified in (1). In order to maintain the attention of the participants the hidden objects varied over trials (cheese, puppet, cookie and ball). During the presentation of the two last context sentences a thought bubble containing the mentioned object was displayed above the mouse's head to make the movie more interesting for the participants (see Figure 7.1/2).

- (1) *Schau mal, vor dem Haus der Maus stehen eine grüne und eine rote Kiste.  
In einer der Kisten ist eine Puppe versteckt.  
Die Maus möchte mit der Puppe spielen.*  
'Look! There is a green and a red box in front of the house of the mouse.  
There is a puppet in one of the boxes.  
The mouse wants to play with the puppet.'

After the context sentences, two test sentences followed in each experimental trial: A complex sentence describing the mouse's belief under the complement taking verb *think* as in (2a) and a simple sentence describing the true state of affairs, which contradicted the content of the complement clause as in (2b). Note that in the experimental trials no hiding procedure was enacted but the information about the mouse's incorrect mental state was given solely verbally. The order of the presentation of the two test sentences were counterbalanced within each child to rule out the strategy to guide the eyes just according to the last mentioned box. To keep the prosody of the two sentence sequence natural they were recorded for each order separately. Prerecorded verbal stimuli were spoken by a female person in a child

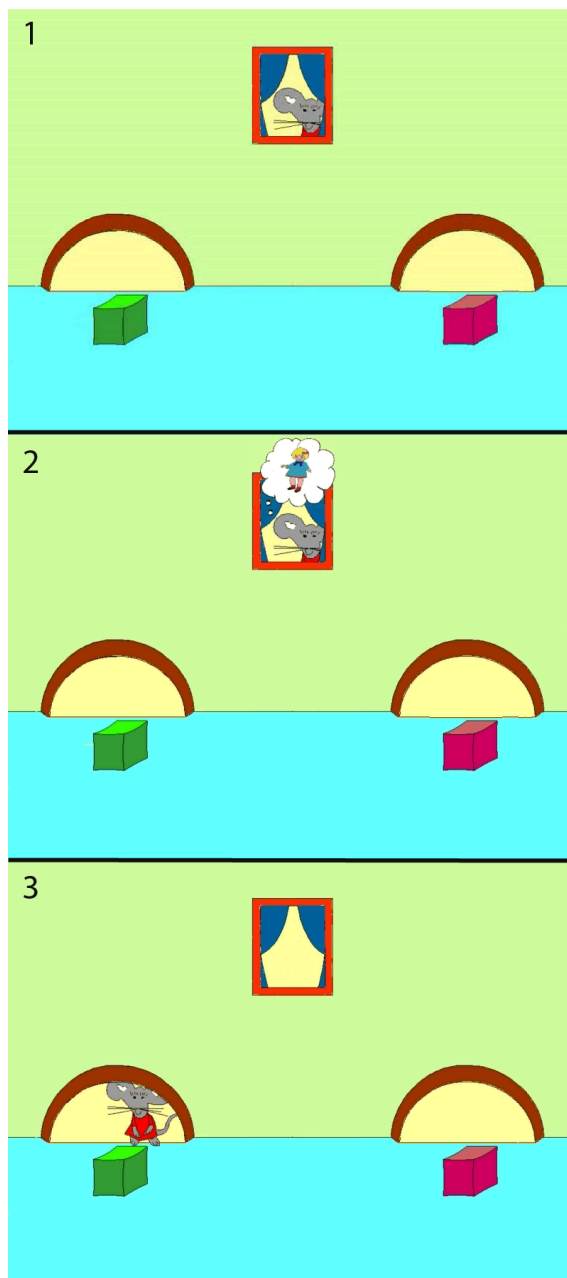


Figure 7.1: Selected frames from Exp. 1: Expressing a false belief

directed manner.

- (2) a. *Die Maus glaubt, dass die Puppe in der grünen Kiste ist.*  
'The mouse thinks that the puppet is in the green box.'
- b. *Die Puppe ist in der roten Kiste.*  
'The puppet is in the red box.'

After the presentation of the test sentences the mouse started to move downstairs, disappeared for 1000 ms and reappeared in the location that corresponded to her belief as expressed in the complement clause (3a) (see Figure 7.1/3). Again, there was no information provided about whether the mouse failed to find the hidden object, i.e. the movie stopped after the mouse reappeared. Between trials, three different animated cartoon figures were presented at the center of the monitor for recapturing or increasing children's attention. These presentations gave way to the next trial as soon as the child focused on the monitor. The entire experiment took about 6 minutes.

### 7.2.3 Apparatus and Procedure

All participants were tested individually in a testing room located at the University. Children's parents were present while testing. A Tobii (Stockholm, Sweden) 1750 Eye Tracker, which is integrated into a 17 TFT monitor and provides eye gaze position data every 20 ms (tracking rate: 50 Hz), was used along with Tobii's ClearView software for stimuli presentation and eye movements recording. Movies were presented in avi-format. Children were seated approximately 60 cm from the monitor in a leaned-back position. Before starting the experiment, a five-point-calibration was performed for each participant (for technical details about the technique and the calibration procedure, see Hofsten et al., 2005). The experiment started when the software determined the calibration to be acceptable. The place of reappearance and the order of the test sentences were counterbalanced across trials within a participant. That is, each child was presented in a random order with two testing trials in which the protagonist reappeared at the left-hand location and two trials in which the mouse reappeared at the right-hand location, in each case the complex sentences preceded the simple sentence in one trial and vice versa in the second one.

### 7.2.4 Data Reduction

Two spatial areas of interests (AOI) were defined (300 x 300 Pixel), each around a mouse hole including the corresponding box in front of it. One AOI is always referred to as *correct AOI* and includes the box which was mentioned in the complement

clause; the other AOI is referred to as *incorrect AOI* and covers the door which is close to that box that was mentioned in the simple sentence. The temporal window was defined to be 1000 ms - that time period the mouse was not visible on the screen.

As dependent measure looking times to each AOI within this second was coded for each participant and trial. To rule out to code random gazes as anticipatory eye movements the minimum of observation duration had to be at least 2 subsequent gazes (approx. 40 ms) within an AOI. The mean proportion of correct anticipations was calculated by dividing the looking time towards that AOI that was expressed in the complement clause by the sum of looking time to both AOIs. Thus, it can range between 0% and 100%, being 50% if there were equal anticipations towards both AOIs, and 100% if there were solely correct anticipations.

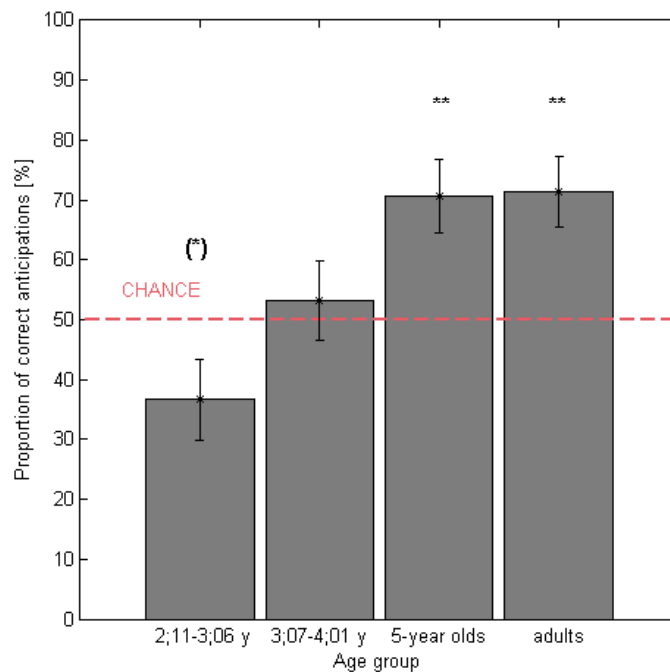
### 7.2.5 Results

To check for effects of the position of the complement sentence within a trial, paired t-tests were performed within each age group comparing mean proportions of correct anticipations in trials in which the complex sentence preceded the simple sentence against trials in which the complement sentence followed the simple sentence. There was no effect of the order of test sentences in all experimental groups (all  $p$ 's > .13). Thus, it can be ruled out that participants just reacted on the last mentioned box. For that reason the following analyses were conducted with averaged data across all four experimental trials. Figure 7.2 shows the mean proportion of correct anticipations in percent for all age groups.

A one-way ANOVA with *age group* (2;11y-3;6y; 3;7y-4;1y, 5-year-olds, adults) as between-subject factor revealed a significant overall-effect of *age* ( $F(3, 71) = 6.401, p = .001$ ). Follow-up t-tests (two-tailed) revealed significant mean score differences between the younger experimental age group (mean score = 36.6%) and each control group (mean score<sub>5-year-olds</sub> = 70.6%,  $t(33) = -3.59, p < .001$ ; mean score<sub>adults</sub> = 71.3%,  $t(36) = -3.79, p < .001$ ). The mean score of the older experimental group (53.2%) differed marginally from the mean score of 5-year olds ( $t(35) = -1.88, p = .068$ ) and significantly from adults ( $t(38) = -2.03, p < .05$ ). No significant group differences were found between experimental groups ( $t(38) = -1.74, p = .09$ ) and between control groups ( $t(33) < 1$ ).

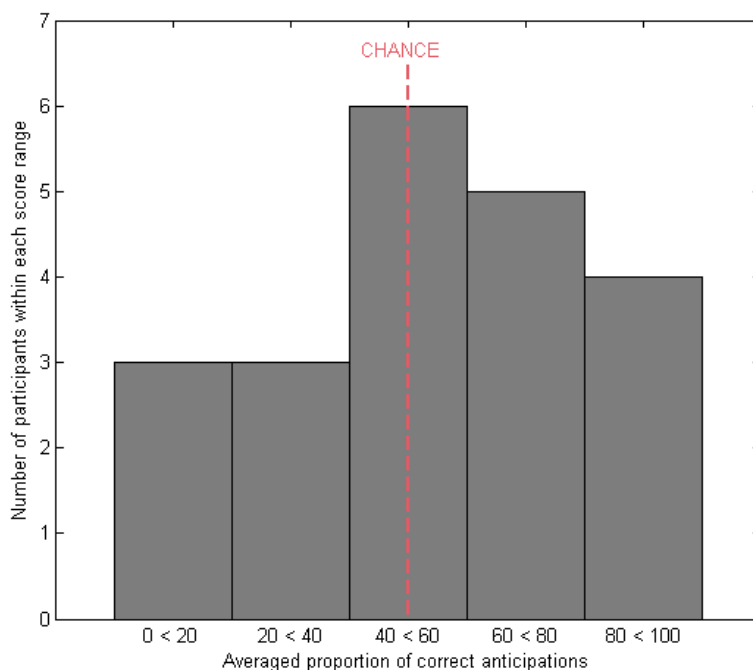
For comparisons to chance level (50%), one-sample t-tests were performed on the data for each participant group. Scores of both control groups were significantly different from chance (5-year olds:  $t(15) = 3.35, p < .01$ ; adults:  $t(18) = 3.60, p < .01$ ). For the 3-year-olds there a two different patterns: the younger experimental group looked marginally below chance ( $t(18) = -1.92, p = .07$ ) and thus show a tendency to anticipate rather towards the incorrect AOI, whereas the mean score of the older experimental children was on chance level ( $t(20) < 1$ ).





**Figure 7.2:** Mean proportion of correct anticipations for all age groups. 100 relate to a time window of one second; error bars indicate standard deviations. *Note.* (\*) $p \leq .07$ , (\*\*) $p \leq .01$ .

To clarify if this averaged observed pattern is a homogeneous pattern within the older experimental group or if this age group consists of a more or less equal amount of children who already look correctly, the distribution of the individual proportion of correct anticipations within the older experimental group was examined (see Figure 7.3). Descriptively speaking, there are participants who showed a quite clear anticipation tendency in either direction. That is, some of them anticipated the reappearance rather in the incorrect area of interest ( $N=3$ ) and almost the same amount show more correct anticipations ( $N=4$ ) on average. That means, it is more a heterogeneous pattern within this older age group. However, this was not correlated with age (in months) ( $r = -0.27, p = .25$ ).



**Figure 7.3:** Frequency distribution of individual averaged proportion of correct anticipations for 3;7- to 4;1-year-old participants.

## 7.2.6 Discussion and Motivation of Experiment 2

The data collected in Experiment 1 suggest first of all that the anticipatory eye movement paradigm can be adopted to measure children's comprehension of complex sentences because both control groups showed the expected looking behavior. Anticipatory eye movements in 3-year-olds, however, did not reveal any early precocious understanding of sentential complements: Gazes of children aged between 2;11 and 3;6 years were apparently guided by the semantic content of the simple non-embedded sentence, whereas gazes of children aged between 3;7 and 4;1 years indicate that they were in a transitional stage because they did not show a clear anticipation tendency. This result is in line with studies using explicit measurements (de Villiers et al., 1990; de Villiers, 1995, 1999; de Villiers and Pyers, 2002), but it is unexpected because an indirect, task-free method was used to decrease production and performance demands. Recent language acquisition studies have shown that the use of eye-tracking provides a more sensitive measure of language processing in young children than traditional off-line comprehension methods. For example, it has been found that children's looking behavior reflects correct interpretation of the focus particle *also* (Höhle et al.,

2009), of pronouns (Sekerina et al., 2004), and of verbal inflection (Brandt-Kobelev and Hohle, 2010).

However, results of Experiment 1 did not reveal any linguistic competence in 3-year-olds suggesting that the comprehension of sentential complementation is limited in children younger than 4 years of age. Two possible explanations are conceivable. First, poor performance of the 3-year-olds might be due to syntactic complexity of complement sentences and children fail to incorporate the embedded clause under the scope of the matrix verb. Second, the task required the children to understand FB. There were two representations of the location of the object given in the context: the believed one and the actual one. To anticipate the reappearance of the mouse correctly one has to understand that someone will act on her/his belief about the object's location, irrespective of its real location. In this respect, the task differs from the memory-for-complement task. Whereas in the memory-for-complement task the child is required to report the content of the person's FB, the current task relies upon children's understanding that a FB leads to a person's mistaken action, that is they have to infer an upcoming action based on the linguistic representation. Thus, performance might be due to conflicting information about reality which is also referred to in the ToM literature as *pull of the real/ reality bias* (Birch and Bloom, 2003; Leslie et al., 2005; Moriguchi et al., 2008; Southgate et al., 2007). As a consequence, it might be not the complexity per se what makes sentential complements hard to comprehend, but its potency to express false propositions.

In order to eliminate the potential confound between the cognitive and linguistic skills, Experiment 2 was conducted in which children's expectations were measured after hearing the content of a person's belief. This time, no information about whether this belief matches with reality or not was given. Thus, no conflicting propositions were presented and children's anticipatory looking behavior is assumed to reflect solely their ability to process complex sentences.

To avoid a simple lexical reaction to the mentioned location in the complement clause the negative particle *nicht* ('not') was added to the matrix clause like in (3a). That is, the incorrect place of reappearance was mentioned in the embedded sentence. The scope of negation is indicated in (3a) by underlining. Thus, the processing of the matrix clause is crucial for the interpretation of the sentence, i.e. children's looking behavior can not be guided solely by the lexical information given in complement clause (green box). Only the correct complement analysis allows *the puppet is in the green box* to be in the scope of negation. To tease apart effects of the comprehension of negation and comprehension of embedded sentences, a first control condition was added. Here the negation occurred in the complement clause as in (3b). In this condition, negation directly applies to *in the green box*, and therefore embedding is not required.

- (3) a. The girl does not think that the puppet is in the green box.  
 b. The girl thinks that the puppet is not in the green box.

If children are proficient in processing negation they are expected to anticipate correctly in the control condition (3b) without necessarily having processed the dependence of the matrix clause. This does not apply to the experimental condition (3a): When a complement analysis is not available, children could only understand (3a) biclausally as illustrated in (4a), where the scope of negation is limited to the first clause. In this case, the children should anticipate the girl to reappear at the green box. For (3b), however, a biclausal understanding amounts to (4b) and therefore still anticipation to the non-green box is predicted.

- (4) a. The girl is not thinking. The doll is in the green box.  
 b. The girl is thinking. The doll is not in the green box.

If children fail to understand negation in general, participants should show poor performances in both, the experimental and the control condition. In both conditions, correct anticipation requires participants to derive from *The doll is not (assumed to be) in the green box* to *The doll may be (assumed to be) in the red box*. If children can assume a doll to be present in the experimental setup, this is a logical inference since there are no alternative places where the doll might be. Nevertheless, children might have difficulty deriving this inference. Therefore, sentential complements without negation as a second control condition were included. Furthermore, children's comprehension of sentential negation in simple sentences was tested in two off-line tasks.

## 7.3 Experiment 2: Expressing a Belief

### 7.3.1 Participants

Forty-one German-acquiring children aged between 3 and 4 years (23 female) were tested. They did not participate in one of the other experiments reported in the present work. This group was split into a younger and older experimental age group. The younger group consisting of 21 children is slightly younger on average than in Experiment 1: 3;2 years (range 3;0–3;5, 10 female); the mean age of the older group is 3;10 (range 3;9–3;11, 13 female) and thus identical with the experimental group of Experiment 1. Again, 5-year-old children were tested (N=16, mean age: 5;1, range 4;11–5;3, 6 female) as well as adults (N=25, mean age: 22 years, range 18–34 years, 22 female) as control groups. According to a questionnaire which was filled in by the parents all children were developing normally and were growing up

in monolingual homes. To check for red/green color blindness children were asked at the beginning to identify a red-colored and green-colored elephant on a painting. This was done because the colors red and green were used in the eye-tacking task. All of the participants were able to discriminate the colors.

## 7.3.2 Material

### 7.3.2.1 Eye-tracking Task

The stimuli consisted of two familiarization trials, followed by nine test trials. Similar to Experiment 1, the visual scene of the experiment contained a house with two doors and a window with a brown-haired girl standing at it (see Figure 7.4/1). In the familiarization trials, the visual scene and the context were introduced: two boxes, a hidden object (which was also displayed in a thought bubble, see Figure 7.4/2) and the fact that the girl wants to obtain this object were mentioned as in (5). In the familiarization trials the real location of the object was mentioned in a simple sentence as in (6). The girl then started to move downstairs, disappeared for 1000 ms and showed up in the door corresponding to the mentioned box and found the hidden object very pleased (she jumps and laughs). While the protagonist was disappearing a chime sounded and both doorframes were illuminated simultaneously. The purpose of the familiarization trials was to teach the participants (a) that the girl in the house wants to obtain the object after she started to move and (b) that she has access to both exits and always takes the shortest way to that box she believes to find the object in. Moreover, the illuminated doorframes and the simultaneous sound should enhance anticipatory looks.

- (5) *Schau mal, vor dem Haus des Mädchens stehen eine rote und eine grüne Kiste.  
In einer Kiste ist ein Keks versteckt.  
Die Maus möchte den Keks essen.*

‘Look! There is a green and a red box in front of the house of the girl.  
There is a cookie in one of the boxes.  
The girl wants to eat the cookie.’

- (6) *Der Keks ist in der roten Kiste.*  
‘The cookie is in the red box.’

In the testing phase, the introductory sentences (5) at the beginning of each trial were limited to the information about the hidden object and the motivation of the girl to obtain it. There were three test conditions: one experimental condition

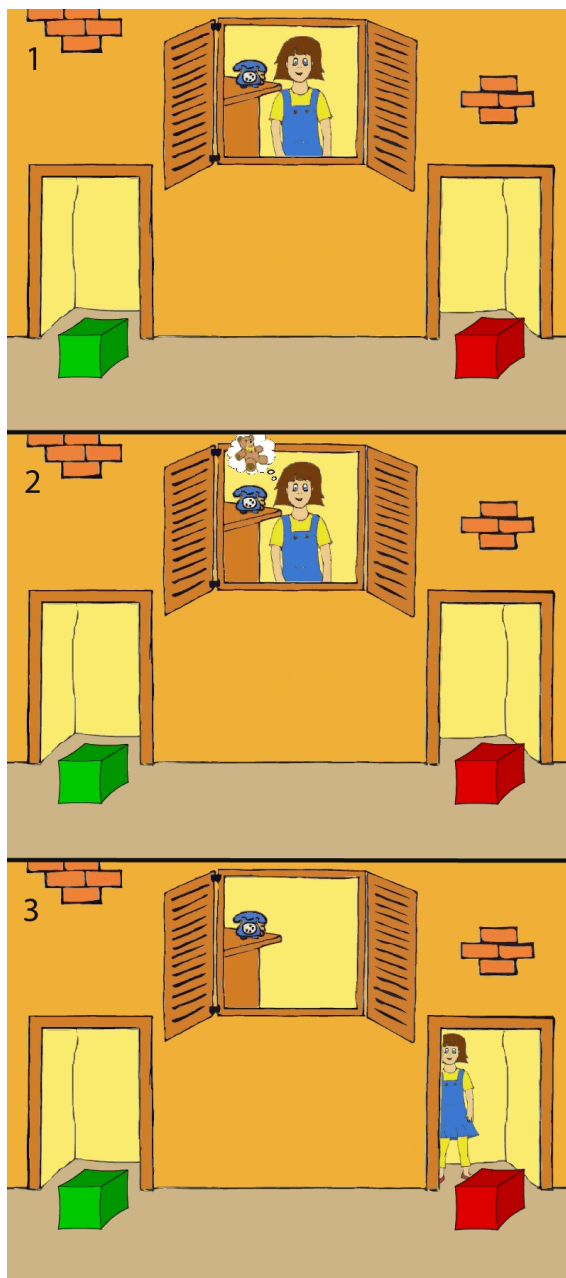


Figure 7.4: Selected frames from Exp. 2: Expressing a belief

and two control conditions in order to account for possible effects originating in the (mis)comprehension of negation. In the *experimental condition*, the negation is realized in the matrix clause as in (7a) and thus scopes over the complement clause. In the control condition *negation*, the prepositional phrase in the complement clause (*in the green box*) is negated since the negative marker is placed in the complement clause as exemplified in (7b). In the control condition *no negation*, no clausal negation was realized (7c). Instead, the adverb *fest* ('strongly') was included in the matrix clause in order to use a similar word order as in the experimental condition.

- (7) a. *Das Mädchen glaubt nicht, dass der Keks in der grünen Kiste ist.*  
'The girl does not think that the cookie is in the green box.'  
b. *Das Mädchen glaubt, dass der Keks nicht in der grünen Kiste ist.*  
'The girl thinks that the cookie is not in the green box.'  
c. *Das Mädchen glaubt fest, dass der Keks in der roten Kiste ist.*  
'The girl strongly thinks that the cookie is in the red box.'

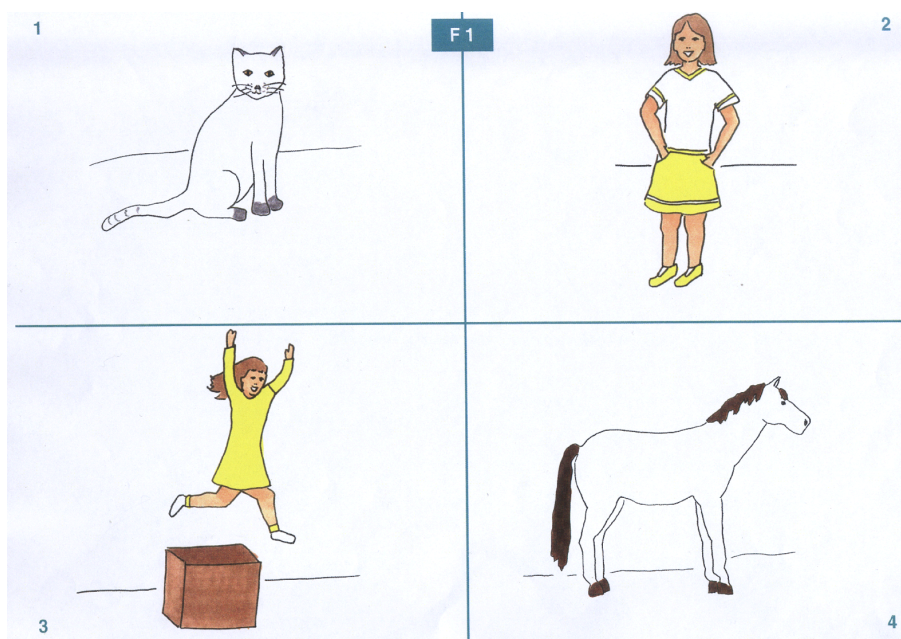
Each condition consisted of three sentences yielding nine test sentences in total. In order to maintain the attention of the children, the "hidden" object varied over trials (ball, duck, teddy, car, book, candy, apple, chocolate, lolly). In contrast to the familiarization phase, the girl upstairs disappeared for 1600ms. Besides, it was not shown in the test trials whether the girl finds the object or not. This was done to avoid giving any information about the truth value of the expressed belief.

### 7.3.2.2 Negation Comprehension Tasks

In order to measure children's comprehension of negation in simple sentences, two tasks were conducted: a sentence-picture matching task and a non-pictorial enactment task.

**Sentence-Picture Matching Task** A subtest (subtest F) was taken from the TROG-D (Test for Reception of Grammar, German version, Subtest U Subordination; Fox, 2006) to assess children's ability to comprehend negated sentences. Participants were shown four pictures per item (as illustrated in Figure 7.5). The child was presented with a simple sentence containing negation like in (8) and was asked then to point to that picture which matches with the sentence (upper right picture in Figure 7.5). The task contained four items in total.

- (8) *Das Mädchen springt nicht.*  
'The girl does not jump.'



**Figure 7.5:** Comprehension of Negation: Example of an item presented in the sentence-picture matching task

There were three types of distractors: one picture illustrated the subject of the sentence and the negated action (e.g., a jumping girl). The two other were lexical distractors displaying different characters (e.g., a horse or a cat), both are not performing the negated action but instead doing something else (e.g. standing or sitting). In order to familiarize children with the task of finding the matching picture four practice items were presented at the beginning of the task. Here the child was required to find the picture which matches to a noun. The visual stimuli were presented in a picture book, sentences were spoken by the experimenter. A list with all items in their presented order appears in Appendix A.4.

**Enactment Task** This task was devised in order to tap children's understanding of negation in a rather natural game play situation without using pictorial stimuli. Furthermore, it tests whether children are able to make an inference from a negated sentence given that the context set an appropriate alternative set. This was done



because in the eye-tracking task, participants have to build up a kind of alternative set to anticipate the girl's upcoming action correctly: If the girl does *not* think that the object is in the red box, then, in turn, it is most likely that she will look for it at the green box.

There were two parts of the task: the *sorting* game and the *finding* game. In the sorting game, two boxes (a blue and a yellow one) and five children figures (two boys and three girls) were introduced to each participant via the following story: *Guck mal. Das hier sind zwei Schachteln, stimmts? Eine blaue und eine gelbe. Die sind jetzt aus Spaß Häuser, einverstanden? Und das hier sind fünf Kinder. Die haben miteinander auf dem Spielplatz gespielt und müssen nun nach Hause gehen, weil es dunkel wird. Jedes Kind wohnt entweder im blauen Haus oder im gelben Haus.* ('Look! There are two boxes, right? A blue one and a yellow one. For fun, we do as if they were houses. And there are five children. They played together at the play yard. Now they have to go home because it is getting dark. Each child is living either in the blue or in the yellow house.'). Children then were asked to help each figure to find its house after they had heard a "hint". These hints comprised two filler trials with simple non-negative sentences like in (9a) and three testing trials with simple negative sentences like in (9b). Sentences were performed live by the experimenter.

- (9) a. *Dieses Mädchen wohnt im blauen Haus.*  
'This girl lives in the blue house.'  
b. *Dieses Mädchen lebt nicht im blauen Haus.*  
'This girl does not live in the blue house.'

In the finding game that also comprised five items, the two boxes were introduced as two potential hiding places. Participants were told that they would play a hiding and finding game in which the experimenter hides an object in one of the two boxes while the child is not watching. The child was then asked to find the hidden object after a sentence ("a hint") was presented. In two filler trials, children were presented with simple non-negative sentences like in (10a); in the remaining three trials, the critical sentences contained the negation marker *not* like in (10b). If children are proficient in processing negative sentences they are expected to choose the non-mentioned box in the testing trials. To increase participants' motivation they were told that they could keep the objects (e.g. a sticker, balloon) in case they found them. Again, sentences were presented live by the experimenter. The number of correct responses could range from 0 to 6 for the complete enactment task; for each subtest from 0 to 3.

- (10) a. *Der Luftballon ist in der blauen Schachtel.*  
'The balloon is in the blue box.'
- b. *Der Luftballon ist nicht in der blauen Schachtel.*  
'The balloon is not in the blue box.'

### 7.3.3 Procedure and Apparatus

Each participant was tested individually in a testing room located at the University. The parents of the children were present while testing, though they were asked to not influence the child's reactions. At the beginning, the eye-tracking experiment was conducted. Each participant was either placed in a leaned-back chair or was sitting at the parent's lap, approximately 60 cm in front of a monitor on which the animated cartoons were presented. Children were instructed to watch cartoons. A Tobii T120 Eye Tracker (Stockholm, Sweden) which is integrated into a 17-in. TFT monitor was used along with the Tobii Studio software package for stimuli presentation and eye movement recording. The Tobii T120 Eye Tracker tracks the position of the eyes 120 times per second. Video stimuli were presented in avi-format. A 5-point calibration was completed before the stimulus presentation.

Two familiarization trials and nine test trials (three per condition) were presented to each participant. Half of the participants started with the familiarization trial where the protagonist finds the object in the red box; the other half started with the other familiarization trial (green box). The nine test trials were presented in a pseudorandomized order, that is there were three blocks each containing one trial per condition. The order of test trials within the blocks changed and the place of reappearance was counterbalanced across items and participants. Between blocks, three different animated cartoon figures were presented at the center of the monitor for recapturing or increasing children's attention. The next block was started as soon as the child focused on the monitor.

After the eye-tracking experiment, the two negation comprehension tasks were conducted with children participants starting with the sentence-picture matching task. During the off-line tasks participants sat at a table suitable for children. Their responses were recorded at the time of testing. The entire experiment took about 20 minutes.

### 7.3.4 Data Reduction

Two spatial areas of interests (AOI) were defined (250 x 350 Pixel), each around a door including the corresponding box in which one AOI is referred to as *correct AOI* and the other one as *incorrect AOI*. As for the experimental condition and the control

condition *negation*, the location expressed in the complement clause corresponds to the box within the incorrect AOI, whereas for the control condition *no negation* the mentioned box relates to the correct AOI. The temporal window was defined to be 1600 ms: that time period the girl just disappeared from the window until she reappeared at one of the doors. As core measure of action anticipation, the looking proportion within each of both AOIs was calculated first for each trial and participant, then averaged across trials within participant and finally averaged across participants. As in Experiment 1, the minimum of observation time within an AOI was set to be 40 ms. The mean proportion of correct anticipations was calculated by dividing the looking time towards the correct AOI by the sum of looking time to the correct and incorrect AOIs. Thus, it can range between 0% and 100%, being 50% if there were equal anticipations towards both AOI, and 100% if there were solely correct anticipations. One-sample t-tests (two-tailed) were performed on the data for each participant group and condition to test whether the mean proportion of correct anticipations differs from a chance (50%). Furthermore, data were entered into a 3 (condition as a within-subject factor) x 4 (age group as a between-subject factor) repeated ANOVA to test for age and condition effects.

### 7.3.5 Results

#### 7.3.5.1 Eye-tracking Task

Mean proportion of correct anticipations and *t*-values and *p*-values of one-sample t-tests (tested against 50%; chance level) are given in Table 7.1 for each participant group and condition.

T-tests revealed that participants of both control groups scored significantly above chance in the experimental condition and control condition *no negation*: There were more looks in anticipation towards the correct AOI than to the incorrect AOI. The mean proportion of correct looks in the control condition *negation* did not reach significance for the 5-year olds but for the adults. Within the 3-year olds a heterogeneous pattern was observed: Whereas the younger experimental group (mean age 3;2 years) did not display a clear anticipation tendency in any of the conditions the older experimental group (mean age 3;10 years) looked significantly longer towards the correct AOI than towards the incorrect AOI in all conditions. These data are illustrated in Figure 7.6.

A repeated measure ANOVA with *age group* as between-subject factor and *condition* as within-subject factor revealed a significant effect of *age group* ( $F(3, 67) = 19.84, p < .001$ ) and a marginal effect of *condition* ( $F(2, 134) = 3.04, p = .056$ ). The interaction between *condition* and *age group* was not significant ( $F(3, 134) < 1$ ). Pair-wise comparisons between age groups (averaged across conditions) showed that the overall mean score of correct anticipations of younger 3-year-olds differed significantly

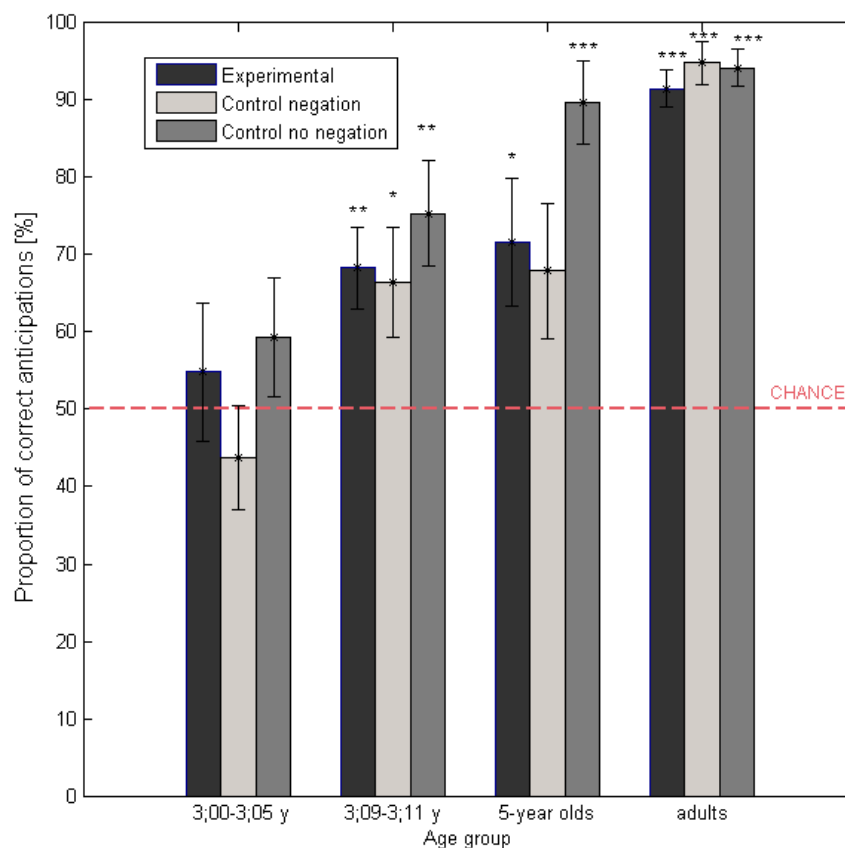
**Table 7.1:** Mean proportion of correct looks and one-sample t-tests scores for all conditions and age groups (standard deviation in brackets).

Condition	Age group	Proportion of correct antizipations	one-sample t-test against chance (50%)
Experimental condition	3;0–3;5 years (N=18)	54.85 [40.85]	$t(17) = 0.50$ $p = 0.62$
	3;9–3;11 years (N=19)	68.31 [23.87]	$t(18) = 3.34$ $p = 0.004$
	5-year-olds (N=15)	71.55 [33.10]	$t(14) = 2.52$ $p = 0.024$
	Adults (N=25)	91.53 [11.71]	$t(24) = 17.74$ $p < 0.001$
Control condition <i>negation</i>	3;0–3;5 years (N=16)	43.70 [30.73]	$t(15) = -0.82$ $p = 0.42$
	3;9–3;11 years (N=19)	66.44 [32.02]	$t(18) = 2.24$ $p = 0.038$
	5-year-olds (N=14)	67.90 [35.25]	$t(13) = 1.9$ $p = 0.09$
	Adults (N=25)	94.82 [13.67]	$t(24) = 16.39$ $p < 0.001$
Control condition <i>no negation</i>	3;0–3;5 years (N=18)	59.34 [35.47]	$t(17) = 1.12$ $p = 0.28$
	3;9–3;11 years (N=18)	75.38 [30.29]	$t(17) = 3.55$ $p = 0.002$
	5-year-olds (N=15)	89.73 [21.52]	$t(14) = 7.15$ $p < 0.001$
	Adults (N=25)	94.21 [12.23]	$t(24) = 18.08$ $p < 0.001$

from those of the older 3-year-olds, 5-year-olds and adults ( $p < .05$ ,  $p < .01$  and  $p < .001$ , respectively), while older 3-year-olds did not differ in their mean score from 5-year-olds ( $p = .30$ ), but from adults ( $p < .001$ ). In addition, 5-year-olds showed a significantly lower mean score of correct anticipations than adults ( $p < .001$ ).

A significant difference across all age groups was found between the experimental condition and control condition *no negation* ( $t(73) = -2.19$ ,  $p < .05$ ) and a marginally significant difference was found between both control conditions ( $t(70) = -1.97$ ,  $p = .052$ ) with a higher score of correct anticipations for the condition *no negation*. The experimental condition and the control condition *negation* did not differ overall ( $p = .59$ ). These condition differences are presumably due to the pattern of the older experimental group and 5-year-olds.

Examination of the age group differences for each condition separately revealed significant differences for the experimental condition between adults and all groups of children (younger 3-year-olds:  $t(41) = -4.27$ ,  $p < .001$ ; older 3-year-olds:

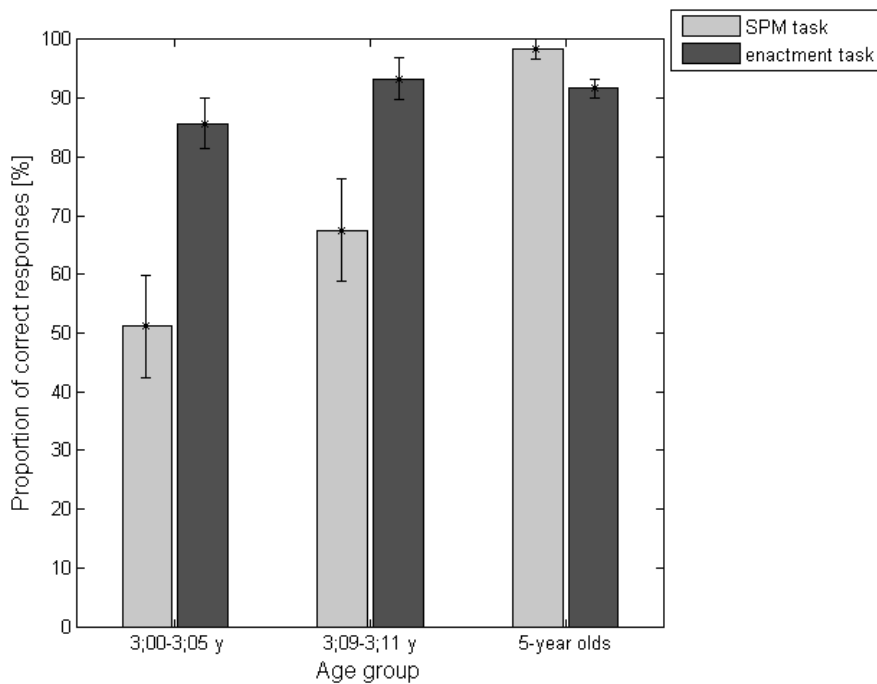


**Figure 7.6:** Mean proportion of correct anticipations in each condition and each age group. Error bars show standard deviations. The red line indicates the 50%-chance-level. *Note.* \* $p \leq .05$ , \*\* $p \leq .01$ , \*\*\* $p \leq .001$ .

$t(42) = -4.25, p < .001$ ; 5-year-olds:  $t(38) = -2.76, p < .01$ ). None of the pairwise comparisons between the three groups of children reached significance within the experimental condition (all  $p$ 's  $> .2$ ). As for the control condition *negation*, the mean score of correct anticipations of younger 3-year-olds differed significantly from those of the older 3-year-olds and adults ( $t(33) = -2.13, p < .05$ ;  $t(42) = -4.25, p < .01$ , respectively), and marginally from 5-year-olds ( $t(42) = -2.01, p = .054$ ). In addition, adults performed significantly better than older 3-year-olds ( $t(42) = -3.99, p < .001$ ) and 5-year-olds ( $t(37) = -3.41, p < .01$ ). Within the control condition *no negation* adults differed significantly from both groups of 3-year-olds (younger 3-year-olds:  $t(41) = -4.57, p < .001$ ; older 3-year-olds:  $t(42) = -2.82, p < .01$ ), but not from 5-year-olds. In addition, younger 3-year-olds showed significantly lower mean scores of correct anticipations compared to 5-year olds ( $t(31) = -2.90, p < .01$ ).

### 7.3.5.2 Negation Comprehension Tasks

Dependent t-tests (two-tailed) revealed that performance did not differ between the two subtests of the enactment task (sorting game and finding game) within each age group (younger 3-year-olds:  $t(20) < 1$ ; older 3-year-olds:  $t(19) = 1, p = .33$ ; 5-year-olds:  $t(14) = 1.47, p = .16$ ). Therefore, the following analyses were conducted with averaged data across all six experimental trials. Figure 7.7 shows the mean proportion of correct responses in the enactment task and the sentence-picture matching task for all three age groups.



**Figure 7.7:** Mean proportion of correct responses in each negation comprehension task and each age group. Error bars show standard errors.

To test whether the performance within each age group differed significantly from chance, one-sample t-tests (two-tailed) were conducted. Given that there is a 25% chance to point to the correct picture in the sentence-picture matching task, all age groups performed significantly better than chance (younger 3-year-olds:  $t(20) = 3.01, p < .01$ ; older 3-year-olds:  $t(19) = 4.88, p < .001$ ; 5-year-olds:  $t(14) = 4.4, p < .001$ ). However, this chance level has to be considered critically.

Since two out of three distractor pictures displayed neither the character associated with the subject of the sentence nor the (negated) action one could consider the chance level to be really 50% instead of 25%. If tested against 50% chance level, younger 3-year-olds displayed no significant above-chance performance anymore ( $t(20) < 1$ ) and older 3-year-olds performed marginally significantly better than chance ( $t(19) = 2.01, p = .059$ ). There was no change in the  $p$ -value for 5-year-olds ( $t(14) = 29, p < .001$ ).

In the enactment task in which one out of two locations had to be selected (i.e. 50% chance level), each age group displayed a highly significant above-chance performance (younger 3-year-olds:  $t(20) = 8.22, p < .001$ ; older 3-year-olds:  $t(19) = 12.37, p < .001$ ; 5-year-olds:  $t(14) = 31.55, p < .001$ ).

A repeated measure ANOVA with *age group* as between-subject factor and *task* as within-subject factor revealed a significant effect of age group ( $F(2, 53) = 10.10, p < .001$ ), of task ( $F(1, 53) = 16.89, p < .001$ ) and a significant interaction of age group and task ( $F(2, 53) = 4.35, p < .05$ ). As for the task effect,  $t$ -tests for dependent samples were conducted for each age group separately. Both younger and older 3-year-olds gave significantly more correct responses in the enactment task than in the sentence-picture matching task (younger 3-year-olds:  $t(20) = 3.54, p < .01$ ; older 3-year-olds:  $t(19) = 3.05, p < .01$ ). 5-year-olds performed equally well in both tasks ( $t(14) < 1$ ). Comparisons between age groups revealed that younger and older 3-year-olds did not differ in their performances in both tasks (enactment task:  $t(39) = 1.36, p = .18$ ; sentence-picture matching task:  $t(39) = 1.32, p = .19$ ). However, 5-year-olds outperformed the younger experimental group significantly in both tasks (enactment task:  $t(34) = 2.62, p < .05$ ; sentence-picture matching task:  $t(34) = 4.52, p < .001$ ), and the older experimental group in the sentence-picture matching task ( $t(33) = 3.02, p < .01$ ), but not in the enactment task ( $t(33) = 1.04, p = .31$ ).

### 7.3.6 Discussion

Experiment 2 tested whether children show precocious comprehension of sentential complementation in their anticipatory looking behavior after being confronted with a complex sentence expressing someone's content of belief. In contrast to Experiment 1, no additional information about the truth value of the embedded proposition was given to circumvent the potential interference with children's FB understanding. A negative was added to the complement sentence in the matrix clause to make the processing of the matrix clause crucial for the interpretation of the sentence. First, children's performance in two off-line tasks tapping the understanding of negation in simple sentences will be interpreted before turning to the eye movement data.

**Comprehension of negation in simple sentences** From the data of the enactment task it is evident that 5-year-olds and already 3-year-olds comprehend negation in simple sentences. Participants selected significantly more often the second, unmentioned container that was present in the experimental setup. Thus, they were able to compute the inference *The object is in box B.* from the sentence *The object is not in box A.* This is important with respect to the eye-tracking task because participants had to derive the similar inference when anticipating the upcoming behavior of the protagonist.

However, younger 3-year-olds failed to demonstrate their comprehension ability on negation when tested with the sentence-picture matching task. It is important to note that this holds only when the rigorous criterion (50% chance level) is applied<sup>1</sup>. In addition, also older 3-year-olds showed a better performance in the enactment task than in the sentence-picture matching task. One possible explanation for this difference between tasks is that the sentence-picture matching task might be more demanding than the enactment task because children were presented not only with the linguistic but also with the visual information. It is plausible that the sentence-picture matching task involved additional cognitive processes like for instance inhibitory control as the child was required to simultaneously reject the non-target picture (see also Karmiloff-Smith et al., 1998). This does not apply to the enactment task.

In sum, 3-year-old children are proficient in interpreting simple sentences containing negation when an appropriate context is provided. This is supported by previous studies: In a study by de Villiers and Flusberg (1975), children from age 2 years demonstrated a correct use of negation. In addition, German-speaking children were found to start to produce the negation marker *nicht* ('not') around their second birthday (Wode, 1977).

**Comprehension of complex sentences** As in Experiment 1, adults showed the expected gaze pattern after having heard the test sentences containing a complement clause in all three conditions: In sentences without negation they expected the story's protagonist to reappear in that location she believed to find the object. Respectively, in sentences with negation, they did not expect the girl's reappearance in that location she did not think to be the object's location, but in the alternative place. 5-year-olds also displayed a target-like interpretation of the complex sentences in two out of three conditions with respect to chance-level performance: whereas they showed correct anticipations after sentences without negation and sentences where the complete embedded proposition is negated, 5-year olds had slightly more difficulties if the negation occurred in the complement clause and thus applied to

---

<sup>1</sup>Although the child was required to select one out of four pictures only two of them were considered as actual competitors: the incorrect one showing the negated action and the target picture showing the sentence's subject doing something else.



the prepositional phrase. However, since this is only a marginal difference within the 5-year-olds, eye-tracking still seems to be a very suitable method to measure comprehension of complex sentences. Comparisons between adults and 5-year-olds show that 5-year-olds interpret complex sentences without negation already in an adult-like fashion, whereas performances of the 5-year olds did not reach the performance-level of adults in sentences with negation. This difference is presumably due to the fact that 5-year-olds still have more difficulties in deriving the logical inference which had to be done in negated sentences.

Younger 3-year-olds (mean age 3;2 years) did not show a clear anticipation preference in any of the tested conditions. Hence, they still have difficulties in processing sentences containing an embedded clause when anticipating an upcoming event based on the given information. It can be ruled out that they understood the complex sentence bi-clausally (e.g. The girl is [not] thinking. The object is [not] in the green box.) because this would result in a clear tendency towards the correct place of reappearance in both control conditions (where no embedding is required), and a tendency towards the incorrect place of reappearance in the experimental condition is expected. In addition, apparently children did not apply a simple strategy based on lexical information like 'looking-always-towards-the-mentioned-box'-strategy since this would have led them to anticipate correctly only in the condition where no negation occurred.

Furthermore, comparisons between sentences containing negation and sentences without negation suggest that their difficulties did not solely originate in their comprehension of negation in complex sentences. Instead, it is assumedly a matter of a general effect of the syntactical complexity what make young 3-year-olds still struggle with the interpretation of complement sentences yielding in a rather random looking behavior. Yet, younger 3-year-olds do not interpret the proposition of the complement clause as a simple SVO-sentence which indicates that they are sensitive to the complexity of the sentence. This is probably due to the additional marking by the verb's final position in complement clauses which had been found to be produced from the age of 2;0 years on in German-acquiring children correctly (Rothweiler, 1993; Brandt et al., 2010). Weissenborn et al. (1998) also found that 2-year-olds are sensible to word order violations in subordinate clauses. However, young 3-year-olds cannot interpret complement sentences in a target-like fashion yet. The poor performance of younger 3-year-olds could also originate in their difficulties to compute inferences as it must be concluded from the utterance that the protagonist will choose the not-mentioned box to search for the object. However, in one of the two offline tasks (enactment task) even young 3-year-olds were able to act on the implied utterance (e.g., *The sticker is NOT in the blue box.*) and choose the non-mentioned box to look for the object.

Children between 3;9 and 3;11 years showed a higher amount of correct anticipatory looks in each of the three conditions with respect to chance level. Besides, this performance pattern did not differ from the performance of 5-year olds. This suggests that children from the age of 3;9 years on are proficient in processing complement sentences in a target-like fashion. Still, older 3-year-olds did not perform on an adult-like level yet: In all three conditions the proportion of correct anticipations scored lower in 3;9- to 3;11-year-olds compared to adults.

Yet, further considerations are necessary in order to rule out alternative explanations of the results found in the current experiment. The control condition *negation* (i.e. in which the negative occurred in the complement clause) was designed to isolate the effect of negation from syntactic complexity. Since in this condition the scope of the negation applied only to the prepositional phrase in the complement clause no complement analysis was required to anticipate correctly in the control condition whereas it was crucial in the experimental condition in which the complement clause as a whole was in the scope of negation. Previous findings on children's comprehension of negation in complex sentences indicate that the scope of negation might be not understood correctly until school age (Phinney, 1981; Singer, 1987). More specifically, it was found that children up to 12 years "consistently interpret embedded negatives as matrix negatives" (Singer, 1987, p.195). Phinney (1981) proposed that children apply a S-attachment strategy, that is they tend to attach the negative marker to the verb in the matrix clause irrespective of its position in the sentence. This phenomenon is also referred to as *neg-raising*. The question arises if and to what extent the interpretations of the current results are affected by these findings.

Assuming that children in the current study employed such an attachment rule this would only explain their performance in the control condition but not their performance in the experimental condition. In addition, in both studies cited above the complement-taking verbs used were always factive verbs (e.g., *know*, *surprise*) where the truth value of the embedded proposition does not change regardless of whether the matrix verb is negated or not. If participants interpreted the nonfactive verb used in Experiment 2 (*glauben*, 'think') as a factive verb like in *The girl does not know that the puppet is in the green box.*, this would have led them to anticipate incorrectly in the experimental, but not in the control condition *negation* (i.e. *The girl knows that the puppet is not in the green box.*). Therefore, it can be ruled out that the correct looking behavior in older 3-year-olds and 5-year-olds originate in their misrepresentation of the negative's scope and/or the semantics of the matrix verb.

## 7.4 General Discussion

*Study 2* was set up to examine the onset of sentential complement understanding in tasks with reduced task demands by measuring spontaneous anticipatory eye gaze instead of posing questions. Results of both experiments do not provide evidence that indirect measures reveal a considerably earlier understanding of complement sentences. In Experiment 2, children seem to successfully understand complement clauses from the age of 3;9 years on when the truth value of the embedded proposition is not given. This indicates that poor performance in direct off-line tasks is not due to the requirement of making a conscious decision or tapping into meta-linguistic knowledge which could have masked performance in previous experiments. Rather, 3-year-old children seem to be limited in their interpretation of sentences taking a finite clause as its object under the matrix verb *think*. However, it is important to note that both eye-tracking tasks required an understanding that beliefs (regardless of whether they are true or false) cause behavior, which is not a trivial task (see Howlin et al., 1999; Wellman and Liu, 2004, for models on the development of an understanding of mind). That is, inferring an action from linguistic information involves an additional computational step which is not required in traditional methods relying on direct question. Especially young 3-year-olds' poor performance could originate in their difficulty to infer an action from a mental state.

Yet, there was a difference between the performance pattern between Experiment 1 and Experiment 2 which is worth discussing. While in Experiment 1 even 3;6- to 4;1-year-olds had difficulties in processing sentences expressing a FB when anticipating an upcoming event, in Experiment 2 older 3-year-olds were able to comprehend complex sentences expressing a belief without having the information about the truth value of the embedded proposition. In addition, also in younger 3-year-olds this performance boost was observed: whereas they seem to ignore the complex sentence when inferring the protagonist's future actions in Experiment 1, their looking behavior in Experiment 2 indicates that they are somehow undecided in their expectation. This increase of correct looking behavior is assumedly due to the absence of the conflicting reality information in Experiment 2. If there are not two representations of the object's location (namely the real location and the believed one as in Experiment 1) children are less misled in their expectations. Hence, it is proposed that scope of negation is the more suitable approach when measuring children's ability to process sentential complements.

However, there are several modifications between the two experiments which might have an impact on the outcome and provide alternative explanations for the observed discrepancy. First, the anticipation phase was extended in Experiment 2 from 1000 ms to 1600 ms. Analyses restricted to the first 1000ms in Experiment 2, however, revealed no fundamental changes in the results, especially with regard to the experimental condition. Hence, it can be excluded that the extended anticipation

phase caused the increase of anticipatory looks towards the correct area of interest. Second, the familiarization phase was verbal in Experiment 2, but nonverbal in Experiment 1. It remains possible that in Experiment 1 the link between the preceding (especially FB) context and its particular outcome was not obvious for the younger children. Third, using a non-human protagonist (the mouse) in Experiment 1 might also have an impact on the results. However, as Surian et al. (2007) showed children seem to attribute mental contents not only to humans, but also to non-human agents.

# Chapter 8

## General Conclusions

In this chapter it will be discussed how the findings of the current work contribute to the understanding of the relationship between the comprehension of sentential complements and FB understanding. Both abilities have been assessed with implicit and explicit measures. The chapter is subdivided into three parts: First, a summary of the main findings is given with respect to the research questions raised in the introduction. The second part deals with methodological considerations on the influence of task characteristics on performance in (a) implicit FB tasks and (b) sentential complementation tasks. More specifically, 3-year-olds' poor performance on the nonverbal implicit FB task (*Study 1*) is discussed in the light of previous research that demonstrated an impact of the presentation mode (verbal vs. visual) on children's performance in explicit FB tasks (section 8.2.1). In 8.2.2, the impact of the reality information on children's performance in sentential complementation tasks is discussed. The third part concerns theoretical considerations on the developmental link between the mastery of sentential complements and ToM.

### 8.1 Summary

The objectives of the present work were subdivided into answering three questions as outlined in Chapter 1.

1. Is there a correlation between implicit false belief understanding and the comprehension of sentential complements?

The first question can be partly answered by looking at the findings of *Study 1*. There was no significant interrelation between children's anticipatory eye gazes in the implicit FB task and their performance on sentential complementation tasks although both measures were associated with explicit FB task performance. This is

in line with results by Low (2010, Exp. 1-3) showing that predictive looking was not correlated with complement mastery but with explicit FB understanding. Notably, in the present work no implicit FB understanding was found within the group of 3-year-olds and therefore previous findings on a precocious reasoning ability could not be replicated (Clements and Perner, 1994; Southgate et al., 2007; Ruffman et al., 2001a) (see 8.2.1 for a discussion of possible reasons). Hence, it remains an open question whether the null result is due to the validity of the implicit FB task or to the real absence of a correlation between implicit FB understanding and complement mastery. It is conceivable that the present implicit FB task tapped additionally into the ability to construct the FB story merely by visually given information. That might have enhanced task demands compared to predictive looking tasks that entail besides visual also verbal information (Clements and Perner, 1994; Low, 2010; Ruffman et al., 2001a).

2. Do children show an earlier/precocious comprehension of sentential complements, if methods are used similar to those that have been used to elicit implicit ToM, specifically eye-tracking?

Results of *Study 2* do not provide evidence that implicit measures reveal a considerably earlier understanding of sentential complementation than tasks that require a conscious, verbal reaction from participants. In both experiments of *Study 2* children's expectations about the protagonist's future behavior were measured after being presented either with a complement sentence expressing the protagonist's FB (Exp. 1) or with a complex sentence expressing the protagonist's belief (Exp. 2). Three-year-olds did not display a comprehension of complex sentences (conveyed by their anticipatory looks) if these embedded a false proposition. Considering the performance of Experiment 2, children seem to successfully comprehend complement sentences from the age of 3;9 years on when no additional information about the truth value of the embedded proposition is given. Hence, the results support previous off-line results that complement clauses are acquired only around age 4 (de Villiers, 1995, 1999; de Villiers and Pyers, 2002). This suggests that poor performance in direct off-line tasks like the memory-for-complement task - in which children are asked to report the embedded clause - is not due to the requirement of making a conscious decision which might have been a potential for masking performances in former experiments. Instead, younger 3-year-old children seem to be limited in their comprehension of sentences that take a clause with a finite verb as its complement. This was an unexpected finding because recent language acquisition studies have shown that the use of an indirect, task-free method provides a more sensitive measure of language processing in young children than traditional off-line comprehension methods (e.g., Brandt-Kobele and Höhle, 2010; Höhle et al., 2009; Sekerina et al., 2004).

However, the experimental setup of the tasks of *Study 2* needs to be considered for alternative explanations of the results as well. Both eye-tracking tasks required an understanding that behavior is dependent on beliefs. That is, since children had to infer the protagonist's future action from a verbal description about his/her belief these tasks entail a prediction element that is characteristic of FB tasks. In order to rule out a possible interference with belief understanding, the development of further implicit measures that do not rely on anticipations could open a way to investigate whether preschoolers show any precocious understanding of complement sentences. For example, in a preferential looking paradigm in which two pictures (a matching and a non-matching one) along with the verbal stimuli are presented, a preference for the picture that displays the whole complement sentence would be expected if children are sensitive to the complex syntactic structure.

### 3. Are syntactic properties of sentential complements alone sufficient for the relational link to FB understanding?

Two tasks in *Study 1* were designed to tap children's comprehension of sentential complements without relying on the ability to understand counterfactive propositions which is necessary for solving the broadly used memory-for-complement task. That is, children were not required to reproduce false embedded propositions - a task feature which has been claimed to tap into children's ToM (Adler, 2002; Ruffman et al., 2003) or into children's understanding of the semantic property of sentential complements (Hale and Tager-Flusberg, 2003; Tager-Flusberg and Joseph, 2005), respectively. Instead, participants were asked to select a picture which depicted the whole complement sentence (picture-sentence matching task) and to decide whether a complement sentence described the presented picture correctly (truth-value-judgment task). That is, these tasks involved no counterfactive complement clauses.

Both tasks correlated positively with explicit FB task performance. In addition, a hierarchical logistic regression analysis revealed that this correlation is not simply an age effect but rather complement mastery mediates the interrelation between age and explicit ToM performance. This indicates that understanding the syntactic structure of complement sentences alone is sufficient for FB understanding. de Villiers (1999, 2005a) has claimed that the acquisition of a Point of View (PoV) marker in the child's grammar is crucial for the ToM development as it allows for representing embedded propositions as being false with respect to reality. The two current tasks measured children's ability to interpret the complement clause as being relative to the subject of the matrix clause and not relative to themselves or the speaker which is in turn the understanding of PoV. Hence, the result suggests that the ToM development might be triggered by the ability to comprehend that a subordinate proposition is dependent on the the main clause and not by the comprehension that

an embedded proposition can be false.

However, this conclusion has to be relativized. Due to the small sample size, the logistic regression analysis did not include general language measures as control variables and therefore differs from analyses in previous studies (Low, 2010; de Villiers and Pyers, 2002; Tager-Flusberg and Joseph, 2005). Hence, it can not be ruled out that children's general syntactic development and not specifically the knowledge of sentential complementation is responsible for the effect.

## 8.2 Methodological Considerations

### 8.2.1 Implicit FB understanding: What elicits it?

Unexpectedly, 3-year-olds' visual orienting behavior while watching *nonverbal* FB (and TB) stories did not reveal any implicit FB understanding. They looked to the place where the object really was placed irrespective of the protagonist's belief. This is in line with findings on traditional ToM tasks measuring children's explicit predictions (cf. Wellman et al., 2001). In fact, children's performance on the explicit FB task was associated with their predictive looking behavior: children who failed to exhibit an explicit FB understanding looked rather incorrectly compared to children who passed the explicit FB task.

This result pattern suggests that both the explicit and the implicit FB task measured the same ability and thus there is no gap between indirectly and directly tapped knowledge. However, in the light of previous findings indicating a considerably earlier implicit competence, this conclusion seems premature. The question arises how children come to construct expectations about an actor's behavior based on his FB. The setup of the predictive looking task can be considered as a potential source of explanation (see 6.7 for a detailed discussion). Since the task was as complex as a classic unseen-displacement task (i.e. the object was in the scene while the participant anticipated) and involved no language, it seems conceivable that children are responsive to other's beliefs implicitly only under constrained conditions. That is, to elicit implicit FB understanding either the complexity of the task might have to be reduced (i.e., removing the target object as in Southgate et al.'s (2007) paradigm) or, when keeping the complexity, verbal comments on the ongoing events of the FB story and/or a prompt question prior to anticipation might have to be included to guide children's attention to the critical aspects of the story and/or to the actor's intention (as in Clements and Perner, 1994; Low, 2010; Ruffman et al., 2001a).

There is evidence in the ToM literature that both aspects (complexity and verbal input) affect children's responses in explicit FB tasks. If the target object is not present at the time the FB question is asked children are more likely to predict correctly (Koos et al., 1997; Wellman et al., 2001; Wimmer and Perner, 1983).



Concerning the impact of the presentation mode, it was found that presenting the task solely visually without verbal input makes the explicit FB task harder compared to the standard version (Norris and Millan, 1991, cited by Plaut and Karmiloff-Smith, 1993). In addition, there is evidence that presenting the FB story only verbally, i.e. the participant does not see but is only told about the object transfer, makes the task easier compared to the standard version (Johnson and Maratsos, 1977, pretest; Zaitchik, 1991). Plaut and Karmiloff-Smith (1993) therefore claim that language “provides particularly effective ‘scaffolding’ for symbolic representations” (p.70) and thus helps the child to construct expectations about the actor’s upcoming behavior. A purely verbal presentation of a FB task might additionally make the reality information less compelling because the child does not perceive the actual state of affairs by themselves but is only told about it. The current finding suggests that an impact of the presentation mode might also apply to implicit FB tasks. That is, an additional verbal description of the ongoing events could help children to structure the visually given context and/or the verbal prompt might focus their attention to the protagonist’s goal which then affects children’s anticipatory gazes. In other words, it is possible that a narrative increases children’s involvement in the task by guiding the child’s attention to the critical events of the story and that a verbal prompt helps young children to shift their attention from their own knowledge about the object’s location to the protagonist’s belief. A first step to examine this assumption is to include a verbal narrative into the present implicit FB task and test whether this modification enhances predictive looking performance. An influence of an additional verbal presentation in turn would suggest that the implicit ToM system is pervious to verbal input.

Yet, the result of Experiment 1 of *Study 2* suggests that the scope of the assumption about the influence of verbal input needs to be confined. Here, a stripped-down, verbal-only version of a FB task was used in that it presented a complement sentence expressing the protagonist’s FB and a simple sentence expressing the real state of affairs. Crucially, in the movie it was not shown or narrated what caused the FB (i.e. the unseen displacement). The explicit statement of the protagonist’s FB did not promote 3-year-olds’ anticipatory looking suggesting that the presentation of the whole FB story (i.e. the situational setting) is essential to elicit implicit FB understanding. Hence, verbal input might play a role in terms of assisting the child to grasp and represent the story’s content.

### 8.2.2 Impact of reality information in complementation tasks

In both the explicit and the implicit complementation tasks, additional information about the real state of affairs had a negative impact on children’s performance. That is, if the embedded proposition in the complement clause was false (*Study 1*: reporting-the-complement task, 2-representation condition; *Study 2*: Experiment 1) older

3-year-old children performed at chance. In contrast, children at this age displayed a target-like interpretation of complex sentences when the embedded proposition matched with reality (Study 1: reporting-the-complement task, 1-representation condition) or when no reality information was given (Study 2: Experiment 2). This finding is in line with de Villiers' (2005a) proposed developmental path of the acquisition of sentential complementation. That is, children first understand that mental state and communication verbs embed propositions, and later they understand that these propositions do not have to be true.

However, it is worth asking which process(es) or capacities subserve(s) children's ability to report the false embedded proposition in the memory-for-(false)-complement task. Siegal and Peterson (1996) reported that 3-year-old children were able to judge whether a character is telling the truth, is making a mistake or is lying based on the character's belief or intention. Thus, children's poor performance in the memory-for-complement task is assumedly not caused by a lack of understanding of falsity. This suggests that in the memory-for-complement task children might have difficulties to separate reality from another person's false representation of reality. If we define ToM as the ability to understand that a representation of the situation is different from reality, 3-year-olds might not be able to solve the memory-for-complement task because of a lack in their ToM. It is interesting that de Villiers and Pyers (2002) reported that children's ToM task scores predicted their later memory-for-complement scores. This suggests that reporting false utterances/thoughts relies on children's understanding of FB and would, in turn, support the hypothesis that the memory-for-complement task involves FB understanding. Schulz and Ludwig (2008) reported that FB task passers performed better than FB task failers on interpreting propositions in complement clauses following nonfactive mental state verbs as being undetermined in their truth value. This, again, points towards an impact of children's FB understanding on their comprehension of the nonfactivity status of mental state or communicative complement-taking verbs.

Recapitulating these findings, it appears that tasks tapping the comprehension of complement constructions as the memory-for(-false)-complement task is prone to measure an ability quite similar to that which is measured in ToM tasks. To my knowledge, there is only one study (Tager-Flusberg and Joseph, 2005) that included also true complements in order to tease apart the comprehension of sentential complements from the ability to handle false representations of reality. The present work took a further step towards filling the gap in the literature by designing complement tasks that circumvent a potential confounding influence of ToM abilities.

A related explanation for 3-year-old children's difficulties with the memory-for-complement task comes from the ToM research. Here, the phenomenon of answering in accordance to reality is known as *reality bias* (Mitchell, 1994; Birch and Bloom, 2003; Leslie et al., 2005; Moriguchi et al., 2008). It is conceivable that this bias might also play a role in the memory-for-complement task as (like in

FB tasks) a successful performance on the task requires the child to suppress the (irrelevant) reality information<sup>1</sup> which is presented in a simple sentence. Further research, however, is necessary to examine whether this assumption holds true. One possibility to weaken the reality bias is to make the experimental context of the memory-for-complement task more meaningful in terms of highlighting the potential presence of false utterances. For example, the task could be framed in a kind of ‘detective’ story. That is, the child could be told that s/he is a detective and her/his task is it to identify lying protagonists (and what they said) before presenting the test sentences. In doing so, the content of the complement sentences presumably gains more importance and answering with reality should be less compelling.

### 8.3 Is Implicit ToM independent of syntactic knowledge?

Low (2010) proposed that the linguistic system underpins explicit ToM whereas implicit FB understanding is independent of linguistic representation. Findings of the current work support this assumption<sup>2</sup> in that complement mastery was found to coincide in time with successful explicit ToM task performance but not with children’s spontaneous anticipations in the implicit FB task (*Study 1*, but see 8.2.1 for a methodological consideration concerning the implicit FB task). This indicates that the linguistic-determinism hypothesis (de Villiers, 2000; de Villiers and de Villiers, 2000; de Villiers and Pyers, 2002; de Villiers and de Villiers, 2003; de Villiers, 2005a, 2007) holds only for FB tasks in which a conscious decision is required, that is, whenever the child has to reflect on propositional attitudes explicitly.

It is important to note that the explicit FB task used in *Study 1* (the Smarties task) is highly language demanding as the test question involves a sentential complement (*What does mummy/ daddy think is inside the tube?*). This confound might explain the correlation. However, Low (2010) reported that the correlational link between explicit FB task responses and sentential complement understanding holds for both the nonverbal/ low-verbal and the verbal FB tasks. This supports the claim that the crucial aspect of explicit FB tasks is the requirement of making a direct prediction based on a protagonist’s FB.

Moreover, the results of *Study 2* (Exp. 1) are insightful regarding the question whether verbal information about someone’s FB expressed in a complement sentence can enter the implicit ToM system. 3-year-old children did not display any implicit FB understanding in their anticipatory eye movements after being verbally presented

<sup>1</sup>This ability is also termed *inhibitory control* (e.g., Carlson and Moses, 2001)

<sup>2</sup>Note that in this section the linguistic representation in question is mainly a structural one, presented in complement sentences. It is not identical with the narration guiding children through the FB story and focusing their attention. This was discussed in section 8.2.1. above.

with the protagonist's FB. This indicates that language comprehension in terms of processing sentential complements does not feed into the system underlying implicit ToM. To put it differently, providing the linguistic representation of the protagonist's mental state did not enable children to anticipate the resulting behavior correctly.

Taken together, these findings provide evidence for the existence of two ToM systems: an explicit one that enables children to reflect on propositional attitudes and to verbalize these insights and an implicit system making on-line unconscious belief ascriptions that are reflected in children's spontaneous looking behavior. While previous research offers compelling support for the hypothesis that language, and more specifically the comprehension of complement sentences, provides the means for representing mental states and operate on them (i.e. compute a behavior prediction), relatively little is known about the processes that are involved in the development of implicit ToM.

de Villiers and de Villiers (2003) already acknowledged that a task measuring children's visual orientation is likely to tap knowledge that "is based on empathy or simulation, of what I myself would do if faced with such a set of circumstances." (de Villiers and de Villiers, 2003, p.336). Therefore, they conclude, language might not play a role in implicit FB tasks. Following Ruffman (2004), implicit understanding of belief is based on the (unconscious) detection of behavioral regularities (see also Perner et al., 2007). This assumption would explain the null result in Experiment 1 of *Study 2*. Here, the FB story was not presented visually (i.e. no unseen displacement was shown) but the FB was expressed solely verbally via a complex sentence (and a simple sentence expressing the reality). Therefore, it is possible that children could not build such a *behavior-outcome-link* and thus failed to react correctly with regard to their spontaneous visual orienting.

Several theorists have provided dual-process models that account for the difference between implicit and explicit ToM task performance. For example, Carruthers (2002, 2007) claims that there are two distinct (but related) cognitive systems: a fast and unconscious one which is responsible for children's spontaneous reactions (System 1). Further, he states that System 1 is not uniquely human but shared with other species. This could explain the correct looking behavior of apes in a nonverbal FB task (Krachun et al., 2009). System 2 which is assumed to be conscious, slow and serial is involved whenever the child is required to make a conscious decision (i.e. in traditional, explicit FB tasks). System 2 is assumed to receive its input from the output that is generated by System 1 and additionally from higher-order systems, i.e. the linguistic system and the executive control system. This proposal is supported by empirical data: Low (2010) reported that explicit FB task performance was predicted by both implicit ToM and cognitive flexibility (which is an aspect of executive control). In addition, Clements et al. (2000) showed that children only benefit from a verbal ToM training when they have already developed an implicit ToM indicating that explicit FB reasoning progressively builds upon implicit knowledge.

Apperly and Butterfill (2009) also assume two cognitive mechanisms that account for different levels of ToM task performance by drawing an analogy with number cognition: One mechanism that is language independent, emerges early in development and explains correct performance in preverbal children, and one that is highly language dependent, develops later and is necessary to represent beliefs precisely and to operate on these representations.

Hutto (2008) draws a more radical distinction by claiming that implicit FB tasks do not tap into the understanding of other's beliefs but rather into the understanding of others intentions that can be evaluated in terms of achieving a goal or failing to achieve a goal. Hutto calls them *intentional attitudes* that are opposed to beliefs. Belief, on the other hand, are *propositional attitudes* which can be judged in terms of their truth/falsity. On his account, children acquire the capacity to represent FB through language. Hutto summarizes the role of language in his *Narrative Practice Hypothesis* as follows:

“Long before they acquire a practical grasp of mentalistic concepts, children are able to navigate the social matrix using embodied skills, interacting with others in ways that require no understanding of propositional attitudes [...]. As their command of language increases, they are able to make use of certain syntactic constructions in particular embedded complement clauses. [...] eventually, after actively participating in conversations and exercising their imaginative abilities appropriately, they come by an understanding of that most important of propositional attitudes, belief.” (Hutto, 2008, pp.141-142)

These dual-process models and corresponding empirical data provide also interesting insight for the theoretical debate about the mechanisms that are responsible for the ToM development. Two main theories have been advanced: according to the *Theory-Theory*, a conceptual competence change takes place during the preschool years; children younger than age 4 years fail on ToM tasks because they cannot reason about FB (e.g., Bartsch and Wellman, 1995; Gopnik and Wellman, 1992, 1994; Perner, 1991; Wellman, 1990). In contrast, *early competence accounts* assume an innate, early competence which is masked by performance factors in traditional FB tasks (e.g., Bloom and German, 2000; German and Leslie, 2000; Leslie, 1987, 1994; Scholl and Leslie, 2001; Lewis and Osborne, 1990; Leslie et al., 2005). Dual-process theories may reconcile both theoretical positions because they provide a framework within both an early sensitivity to other's mental states (as indexed by correct anticipatory looking in less-demanding implicit FB task) and the later emerging ability to reflect on other's FB when inferring a behavior (as indexed by children's decision in explicit ToM tasks) can be explained.

However, the question of whether it is, conceptually speaking, the same core ability

remains an important question that has to be addressed by future research. Especially training studies and longitudinal studies are necessary to exam the course of the development of the implicit ToM and it's interrelation with explicit ToM and language.

One way to explore the potential influence of linguistic processes on FB reasoning is to investigate adults' predictive looking performance while they are performing an unrelated verbal task simultaneously. One could pick up the experimental setting used by Newton and de Villiers (2007) and extend this to an implicit FB task. Newton and de Villiers (2007) employed the dual-task paradigm<sup>3</sup> and tested adults on an explicit nonverbal FB task while they either had to perform a verbal task (verbal shadowing) or a nonverbal task (rhythmus shadowing) at the same time. The participants were disrupted in their responses only when they simultaneously performed the verbal interference task. In other words, the verbal dual-task apparently tied up the linguistic resources that were required to solve the explicit ToM task. This indicates that language is needed for the online computation of mental states when a decision is required. The interesting question, then, is if a verbal task interferes also with anticipatory eye gazes. If the dual-process models hold true and implicit ToM relies on nonlinguistic processing resources, adults should have access to implicit FB understanding during both the verbal and a nonverbal interference task.

Experiments using neuro-imaging methods would strongly contribute to answer the question of whether there are two functionally separate ToM systems. Those studies can reveal whether different brain regions are activated (or not) when different kinds of FB tasks (implicit and explicit) are conducted, and if there is an overlap with other brain regions. If the explicit but not the implicit ToM system is dependent from linguistic resources this should be reflected in the activations of brain areas that are involved in language processing when performing an explicit ToM task but not when performing a predictive looking task.

---

<sup>3</sup>In a dual-task paradigm, participants' single-task performance is compared with performance in dual-task condition(s). If performance in a dual-task condition is decreased compared to the single-task condition it is assumed that both tasks access the same resources. Hence, this paradigm allows to determine whether language resources are needed when performing a FB task.

# Bibliography

- Abbeduto, L., & Rosenberg, S. (1985). Children's knowledge of the presuppositions of know and other cognitive verbs. *Journal of Child Language*, *12*(3), 621–641. Cited on page: 48
- Adler, A. N. (2002). Language acquisition and theory of mind development: Evidence for causality? Unpublished paper. Retrieved 06/21/2006, from <http://www.geocities.com/alycat715/TOM.doc>. Cited on page: 2, 30, 35, 54, 131
- Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states. *Psychological review*, *116*(4), 953–970. Cited on page: 22, 137
- Astington, J. W., & Baird, J. A. (2005a). Representational development and false-belief understanding. In J. W. Astington, & J. A. Baird (Eds.) *Why language matters for theory of mind*, (pp. 163–185). Oxford University Press, USA. Cited on page: 13, 26
- Astington, J. W., & Baird, J. A. (2005b). *Why language matters for theory of mind*. Oxford University Press, USA. Cited on page: 25
- Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology*, *35*(5), 1311–1320. Cited on page: 3, 26, 28
- Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2008). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung (12. vollst. überarb. Aufl.)*. Berlin: Springer. Cited on page: 90
- Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, *14*(3), 110–118. Cited on page: 22
- Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, *20*(3), 191–208. Cited on page: 20
- Baron-Cohen, S. (2000). Theory of mind and autism: A review. *International review of research in mental retardation*, *23*, 169–184. Cited on page: 32
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind". *Cognition*, *21*(1), 37–46. Cited on page: 1, 11, 32, 61

- Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind*. Oxford University Press. Cited on page: 10, 12, 46, 137
- Bickerton, D. (1995). *Language and human behavior*. Seattle: University of Washington Press. Cited on page: 28
- Birch, S. A. J., & Bloom, P. (2003). Children are cursed: An asymmetric bias in mental-state attribution. *Psychological Science*, *14*(3), 283–286. Cited on page: 5, 17, 111, 134
- Bloom, L., Rispoli, M., Gartner, B., & Hafitz, J. (1989). Acquisition of complementation. *Journal of Child Language*, *16*(1), 101–120. Cited on page: 43, 46
- Bloom, P. (2002). *How children learn the meanings of words*. The MIT Press. Cited on page: 2
- Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, *77*(1), 25–31. Cited on page: 10, 12, 25, 137
- Boye, K., & Harder, P. (2007). Complement-taking predicates: Usage and linguistic structure. *Studies in language*, *31*(3), 569–606. Cited on page: 41, 47
- Brandt, S., Lieven, E., & Tomasello, M. (2010). Development of word order in German complement-clause constructions: Effects of input frequencies, lexical items, and discourse function. *Language*, *86*(3), 583–610. Cited on page: 43, 46, 125
- Brandt-Kobele, O.-C., & Höhle, B. (2010). What asymmetries within comprehension reveal about asymmetries between comprehension and production: The case of verb inflection in language acquisition. *Lingua*, *120*, 1910–1925. Cited on page: 63, 111, 130
- Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology*, *18*(6), 906–921. Cited on page: 46
- Brown, E. L., & Bull, R. (2007). Can task modifications influence children's performance on false belief tasks? *European Journal of Developmental Psychology*, *4*(3), 273–292. Cited on page: 12
- Brown, J. R., Donelan-McCal, N., & Dunn, J. (1996). Why talk about mental states? The significance of children's conversations with friends, siblings, and mothers. *Child Development*, *67*(3), 836–849. Cited on page: 26
- Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, *112*(2), 337–342. Cited on page: 15
- Call, J., & Tomasello, M. (1999). A nonverbal false belief task: The performance of children and great apes. *Child Development*, *(2)*, 381–395. Cited on page: 13,



39

- Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, 12(5), 187–192. Cited on page: 1
- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, (pp. 1032–1053). Cited on page: 135
- Carpenter, M., Call, J., & Tomasello, M. (2002). A new false belief test for 36-month-olds. *British Journal of Developmental Psychology*, 20(3), 393–420. Cited on page: 14, 15, 17
- Carruthers, P. (2002). The cognitive functions of language. *Behavioral and Brain Sciences*, 25(6), 657–674. Cited on page: 136
- Carruthers, P. (2007). An architecture for dual reasoning. In J. Evans, & K. Frankish (Eds.) *In Two Minds: Dual Processes and Beyond*. New York, NY: Oxford UP. Cited on page: 136
- Chandler, M., Fritz, A. S., & Hala, S. (1989). Small-scale deceit: Deception as a marker of two-, three-, and four-year-olds' early theories of mind. *Child Development*, 60(6), 1263–1277. Cited on page: 13, 25
- Cheung, H., Chen, H. C., & Yeung, W. (2009). Relations between mental verb and false belief understanding in Cantonese-speaking children. *Journal of Experimental Child Psychology*, 104(2), 141–155. Cited on page: 37
- Cheung, H., Hsuan-Chih, C., Creed, N., Ng, L., Ping, S. W., & Mo, L. (2004). Relative roles of general and complementation language in theory-of-mind development: Evidence from Cantonese and English. *Child Development*, 75(4), 1155–1170. Cited on page: 3, 36
- Clahsen, H. (1988). *Normale und gestörte Kindersprache: linguistische Untersuchungen zum Erwerb von Syntax und Morphologie*. J. Benjamins. Cited on page: 41
- Clements, W., Rustin, C., & McCallum, S. (2000). Promoting the transition from implicit to explicit understanding: A training study of false belief. *Developmental Science*, 3(1), 81–92. Cited on page: 23, 96, 136
- Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, 9(4), 377–395. Cited on page: 3, 10, 14, 15, 16, 17, 53, 59, 79, 80, 96, 130, 132
- Crain, S., & Thornton, R. (1998). *Investigations in universal grammar. A guide to research on the acquisition of syntax and semantics*. MIT Press Cambridge, MA. Cited on page: 84
- Dąbrowska, E., Rowland, C., & Theakston, A. (2009). The acquisition of questions with long-distance dependencies. *Cognitive Linguistics*, 20, 571–597. Cited on page: 49, 51, 99

- de Villiers, J. G. (1995). *Steps in the mastery of sentence complements*. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Indianapolis, IN. Cited on page: 28, 49, 110, 130
- de Villiers, J. G. (1999). On acquiring the structural representations for false complements. In B. Hollebrandse (Ed.) *New perspectives on language acquisition*, (pp. 125–136). GLSA, Amherst, MA: University of Massachusetts Occasional Papers in Linguistics. Cited on page: 27, 44, 45, 49, 50, 99, 110, 130, 131
- de Villiers, J. G. (2000). Language and theory of mind: What are the developmental relationships? In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.) *Understanding other Minds. Perspectives from Developmental Cognitive Neuroscience*, (pp. 83–123). Oxford: Blackwell. Cited on page: 2, 26, 135
- de Villiers, J. G. (2005a). Can language acquisition give children a point of view? In J. W. Astington, & J. A. Baird (Eds.) *Why Language Matters for Theory of Mind*, (pp. 186–219). Oxford Press, USA. Cited on page: 2, 25, 26, 27, 36, 37, 44, 45, 46, 55, 99, 131, 134, 135
- de Villiers, J. G. (2007). The interface of language and theory of mind. *Lingua*, 117(11), 1858–1878. Cited on page: 2, 26, 135
- de Villiers, J. G., & de Villiers, P. (2000). Linguistic determinism and the understanding of false beliefs. In P. Mitchell, & K. Riggs (Eds.) *Children's reasoning and the mind*, (pp. 189–226). Hove, UK: Psychology Press. Cited on page: 2, 26, 27, 69, 135
- de Villiers, J. G., & de Villiers, P. A. (2003). Language for thought: Coming to understand false beliefs. In D. Gentner, & S. Goldin-Meadow (Eds.) *Language in mind: Advances in the study of language and thought*, (pp. 335–384). MIT Press. Cited on page: 2, 23, 26, 28, 38, 135, 136
- de Villiers, J. G., & Flusberg, H. (1975). Some facts one simply cannot deny. *Journal of Child Language*, 2(2), 279–286. Cited on page: 124
- de Villiers, J. G., & Pyers, J. E. (1997). Complementing cognition: The relationship between language and theory of mind. In *Proceedings of the 21st Annual Boston University Conference on Language Development*, (pp. 136–147). Cascadilla Press. Cited on page: 2, 28, 49, 50, 96, 100
- de Villiers, J. G., & Pyers, J. E. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development*, 17(1), 1037–1060. Cited on page: xi, 2, 26, 28, 29, 30, 49, 50, 54, 69, 96, 98, 99, 100, 101, 110, 130, 132, 134, 135
- de Villiers, J. G., Roeper, T., & Vainikka, A. (1990). The acquisition of long-distance rules. In L. Frazier, & J. G. de Villiers (Eds.) *Language processing and language acquisition*, (pp. 257–297). Dordrecht: Academic Publishers. Cited on page: 49, 110

- de Villiers, P. A. (2005b). The role of language in theory-of-mind development: What deaf children tell us. In J. W. Astington, & J. A. Baird (Eds.) *Why language matters for theory of mind*, (pp. 266–297). Oxford University Press, USA. Cited on page: 2, 34
- de Villiers, P. A., Burns, F., & Pearson, B. (2003). The role of language in the theory of mind development of language-impaired children: complementing theories. In *Proceedings of the 27th annual Boston University conference on language development*, vol. 1. Cited on page: 2, 33
- Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences*, 22(5), 735–808. Cited on page: 16
- Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge University Press. Cited on page: 46, 47
- Diessel, H., & Tomasello, M. (2001). The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics*, 12(2), 97–142. Cited on page: 43, 46, 47, 49
- Dilling, H., Mombour, W., & Schmidt, M. H. (2009). *Internationale Klassifikation psychischer Störungen. ICD-10 Kapitel V (F). Klinisch-diagnostische Leitlinien*. Bern: Huber. Cited on page: 32
- Doherty, M., & Perner, J. (1998). Metalinguistic awareness and theory of mind: just two words for the same thing? *Cognitive Development*, 13, 279–306. Cited on page: 37
- Doherty, M. J. (2000). Children's understanding of homonymy: Metalinguistic awareness and false belief. *Journal of Child Language*, 27(2), 367–392. Cited on page: 37
- Doherty, M. J. (2009). *Theory of mind: How children understand others' thoughts and feelings*. Psychology Press. Cited on page: xi, 18
- Dunn, J., & Brophy, M. (2005). Communication, relationships, and individual differences in children's understanding of mind. In J. W. Astington, & J. A. Baird (Eds.) *Why language matters for theory of mind*, (pp. 50–69). Oxford University Press, USA. Cited on page: 26
- Farrant, B. M., Fletcher, J., & Maybery, M. T. (2006). Specific language impairment, theory of mind, and visual perspective taking: Evidence for simulation theory and the developmental role of language. *Child Development*, 77(6), 1842–1853. Cited on page: 33
- Farrar, M. J., Johnson, B., Tompkins, V., Easters, M., Zilisi-Medus, A., & Benigno, J. P. (2009). Language and theory of mind in preschool children with specific language impairment. *Journal of communication disorders*, 42(6), 428–441. Cited on page: 33, 34
- Figueras-Costa, B., & Harris, P. (2001). Theory of mind development in deaf children:

- A nonverbal test of false-belief understanding. *Journal of Deaf Studies and Deaf Education*, 6(2), 92–102. Cited on page: 34
- Fisher, C., Gleitman, H., & Gleitman, L. R. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology*, 23(3), 331–392. Cited on page: 45
- Fitch, W., & Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, 303(5656), 377–380. Cited on page: 1
- Flavell, J., Flavell, E., Green, F., & Moses, L. (1990). Young children's understanding of fact beliefs versus value beliefs. *Child development*, 61(4), 915–928. Cited on page: 30
- Fodor, J. A. (1975). *The language of thought*. Harvard University Press. Cited on page: 26
- Fodor, J. A. (1992). A theory of the child's theory of mind. *Cognition*, 44(3), 283. Cited on page: 25
- Fox, A. V. (2006). *TROG-D. Test zur Überprüfung des Grammatikverständnisses*. Idstein: Schulz-Kirchner Verlag. Cited on page: 68, 71, 115
- Freeman, N. H., Lewis, C., & Doherty, M. J. (1991). Preschoolers' grasp of a desire for knowledge in false-belief prediction: Practical intelligence and verbal report. *British Journal of Developmental Psychology*, 9(2), 139–157. Cited on page: 14
- Frith, U. (1989). Autism and "theory of mind". In C. Gillberg (Ed.) *Diagnosis and Treatment of Autism*, (pp. 33–52). New York: Plenum Press. Cited on page: 32
- Gale, E., de Villiers, P., de Villiers, J., & Pyers, E., J. (1996). Language and theory of mind in oral deaf children. In *Proceedings of the 20th annual Boston University conference on language development*, vol. 1, (pp. 213–224). Cited on page: 34
- Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, 44(1-2), 43–74. Cited on page: 22
- Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences*, 4(2), 59–65. Cited on page: 22
- Garnham, W. A., & Ruffman, T. (2001). Doesn't see, doesn't know: Is anticipatory looking really related to understanding or belief? *Developmental Science*, 4(1), 94–100. Cited on page: 17
- Gelman, R., & Gallistel, C. R. (2004). Language and the origin of numerical concepts. *Science*, 306(5695), 441–443. Cited on page: 22
- German, T. P., & Leslie, A. M. (2000). Attending to and learning about mental states. In P. Mitchell, & K. Riggs (Eds.) *Children's reasoning and the mind*, (pp. 229–252). Cited on page: 10, 12, 137

- Gillott, A., Furniss, F., & Walter, A. (2004). Theory of mind ability in children with specific language impairment. *Child Language Teaching and Therapy*, 20(1), 1–11. Cited on page: 33
- Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, 1(1), 3–55. Cited on page: 41, 45
- Goetz, P. J. (2003). The effects of bilingualism on theory of mind development. *Bilingualism: Language and Cognition*, 6(1), 1–15. Cited on page: 37
- Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, 59, 26–37. Cited on page: 1, 11
- Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind & Language*, 7(1-2), 145–171. Cited on page: 10, 12, 137
- Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. Hirschfeld, & S. Gelman (Eds.) *Domain specificity in cognition and culture*, (pp. 257–293). New York: Cambridge University Press. Cited on page: 10, 12, 137
- Gordon, P. (1996). The truth-value judgment task. In D. McDaniel, C. C. McKee, & H. S. Cairns (Eds.) *Methods for assessing children's syntax*, (pp. 211–231). Cited on page: 67
- Hahn, N. (2009). *False belief understanding and language: Developmental relations*. Ph.D. thesis. Cited on page: 14, 25, 32, 33, 38
- Hale, C. M., & Tager-Flusberg, H. (2003). The influence of language on theory of mind: A training study. *Developmental Science*, 6(3), 346–359. Cited on page: 2, 30, 39, 54, 98, 99, 131
- Happé, F., & Loth, E. (2002). Theory of mind and tracking speakers' intentions. *Mind & Language*, 17(1&2), 24–36. Cited on page: 14, 15
- Happé, F. G. E. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, 66(3), 843–855. Cited on page: 32
- Harris, P. L. (2005). Conversation, pretense, and theory of mind. In J. W. Astington, & J. A. Baird (Eds.) *Why language matters for theory of mind*, (pp. 70–83). Oxford University Press, USA. Cited on page: 31
- Harris, R. J. (1975). Children's comprehension of complex sentences. *Journal of Experimental Child Psychology*, 19(3), 420–433. Cited on page: 44, 48
- Hauser, M. D., Chomsky, N., & Fitch, W. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(5598), 1569–1579. Cited on page: 1, 28
- He, Z., Bolz, M., & Baillargeon, R. (in press). False-belief understanding in 2.5-year-olds: evidence from violation-of-expectation change-of-location and unexpected-contents tasks. *Developmental Science*. Cited on page: 21, 22

- Hofsten, C., Dahlström, E., & Fredriksson, Y. (2005). 12-month-old infants' perception of attention direction in static video images. *Infancy*, 8(3), 217–231. Cited on page: 107
- Höhle, B., Berger, F., Müller, A., Schmitz, M., & Weissenborn, J. (2009). Focus Particles in Children's Language: Production and Comprehension of Auch 'Also' in German Learners from 1 Year to 4 Years of Age. *Language Acquisition*, 16(1), 36–66. Cited on page: 110, 130
- Hooper, J. B., & Thompson, S. A. (1973). On the applicability of root transformations. *Linguistic inquiry*, 4(4), 465–497. Cited on page: 47
- Hopmann, M. R., & Maratsos, M. P. (1978). A developmental study of factivity and negation in complex syntax. *Journal of Child Language*, 5(2), 295–309. Cited on page: 48
- Howlin, P., Baron-Cohen, S., & Hadwin, J. (1999). *Teaching children with autism to mind-read: a practical guide for teachers and parents*. Wiley. Cited on page: 127
- Hughes, C., Adlam, A., Happé, F., Jackson, J., Taylor, A., & Caspi, A. (2000). Good test-retest reliability for standard and advanced false-belief tasks across a wide range of abilities. *Journal of Child Psychology and Psychiatry*, 41(4), 483–490. Cited on page: 101
- Hutto, D. D. (2008). *Folk psychological narratives: The sociocultural basis of understanding reasons*. The MIT Press. Cited on page: 137
- Johnson, C. N., & Maratsos, M. P. (1977). Early comprehension of mental verbs: Think and know. *Child Development*, 48(4), 1743–1747. Cited on page: 14, 48, 133
- Johnson, S. P., Amso, D., & Slemmer, J. A. (2003). Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. *Proceedings of the National Academy of Sciences of the United States of America*, 100(18), 10568–10573. Cited on page: 16
- Kagan, J. (2008). In defense of qualitative changes in development. *Child Development*, 79(6), 1606–1624. Cited on page: 21
- Karmiloff-Smith, A. (1992). *Beyond modularity. A developmental perspective on cognitive science*. Cambridge, MA: MIT Press. Cited on page: 16, 96
- Karmiloff-Smith, A., Tyler, L., Voice, K., Sims, K., Udwin, O., Howlin, P., & Davies, M. (1998). Linguistic dissociations in Williams syndrome: evaluating receptive syntax in on-line and off-line tasks. *Neuropsychologia*, 36(4), 343–351. Cited on page: 124
- Kidd, E., Lieven, E., & Tomasello, M. (2005). The acquisition of complement clause constructions: A sentence repetition study. In *Constructions and acquisition: Proceedings of the 2004 Stanford Child Language Research Forum*, (pp. 50–59).

- Cited on page: 46, 49
- Kiese-Himmel, C. (2005). *AWST- R. Aktiver Wortschatztest für 3- bis 5- jährige Kinder*. Beltz Verlag. Göttingen. Cited on page: 71, 72
- Kiparsky, P., & Kiparsky, C. (1970). Fact. In M. Bierwisch, & K. Heidolph (Eds.) *Progress in Linguistics*, (pp. 143–173). The Hague: Mouton de Gruyter. Cited on page: 27, 44
- Koos, O., Gergely, G., Csibra, G., & Biro, S. (1997). *Why eating Smarties makes you smart? Understanding false belief at the age of 3*. Poster presented at the biennial meeting of the Society for Research in Child Development, Washington, DC. Cited on page: 132
- Kovács, Á. M. (2007). *Enhanced executive functions boost performance in theory of mind tasks: the case of bilinguals*. Paper presented at the SRCDC Biennial Meeting, Boston, 30 March. Cited on page: 37
- Kovács, Á. M. (2009). Early bilingualism enhances mechanisms of false-belief reasoning. *Developmental Science*, 12(1), 48–54. Cited on page: 37
- Krachun, C., Carpenter, M., Call, J., & Tomasello, M. (2009). A competitive nonverbal false belief task for children and apes. *Developmental Science*, 12(4). Cited on page: 3, 136
- Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*, 105(2), 395–438. Cited on page: 22
- Léger, C. (2008). The acquisition of two types of factive complements. In A. Gavarro, & M. J. Freitas (Eds.) *Language Acquisition and Development. Proceedings of Gala 2007*, (pp. 337–347). Cambridge: Cambridge Scholars Publishing. Cited on page: 48
- Leonard, L. B. (2003). Specific language impairment: Characterizing the deficits. In Y. Levy, & J. Schaeffer (Eds.) *Language competence across populations: towards a definition of Specific Language Impairment*, (pp. 209–232). Cited on page: 33
- Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind". *Psychological Review*, 94(4), 412–426. Cited on page: 10, 12, 137
- Leslie, A. M. (1994). Pretending and believing: Issues in the theory of ToMM. *Cognition*, 50(1-3), 211–238. Cited on page: 10, 12, 137
- Leslie, A. M., German, T. P., & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology*, 50(1), 45–85. Cited on page: 10, 17, 22, 111, 134, 137
- Lewis, C., & Osborne, A. (1990). Three-year-olds' problems with false belief: Conceptual deficit or linguistic artifact? *Child Development*, 61(5), 1514–1519. Cited on page: 10, 12, 137

- Limber, J. (1973). The genesis of complex sentences. In T. E. Moore (Ed.) *Cognitive development and the acquisition of language*, (pp. 169–185). Cited on page: 46, 47
- Lohmann, H., & Tomasello, M. (2003). The role of language in the development of false belief understanding: A training study. *Child Development*, 74(4), 1130–1144. Cited on page: 2, 30, 31, 35, 37, 98
- Lohmann, H., Tomasello, M., & Meyer, S. (2005). Linguistic communication and social understanding. In J. W. Astington, & J. A. Baird (Eds.) *Why language matters for theory of mind*, (pp. 245–265). Oxford University Press, USA. Cited on page: 37
- Low, J. (2010). Preschoolers' implicit and explicit false-belief understanding: Relations with complex syntactical mastery. *Child Development*, 81(2), 597–615. Cited on page: 3, 5, 10, 23, 38, 39, 53, 54, 79, 96, 97, 98, 100, 101, 130, 132, 135, 136
- Lunn, J. A. (2003). *Is Call and Tomasello's (1999) nonverbal false belief task a test of precocious mental state understanding?*. Paper presented at the Biennial Meeting of the Society for Research in Child Development. Tampa, FL. Cited on page: 13
- Matsui, T., & Miura, Y. (2008). *Pro-social motive promotes early understanding of false belief*. Nature Publishing Group, available from nature preceding <<http://hdl.handle.net/10101/npre.2008.1695.1>>. Cited on page: 15
- Mayes, L. C., Klin, A., Tercyak Jr, K. P., Cicchetti, D. V., & Cohen, D. J. (1996). Test-retest reliability for false-belief tasks. *Journal of child psychology and psychiatry*, 37(3), 313–319. Cited on page: 101
- McMurray, B., & Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy*, 6(2), 203–229. Cited on page: 55
- Miller, C. A. (2001). False belief understanding in children with specific language impairment. *Journal of Communication Disorders*, 34(1-2), 73–86. Cited on page: 33
- Miller, C. A. (2004). False belief and sentence complement performance in children with specific language impairment. *International Journal of Language & Communication Disorders*, 39(2), 191–213. Cited on page: 33
- Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, 78(2), 622–646. Cited on page: 1, 25
- Mitchell, P. (1994). Realism and early conception of mind: A synthesis of phylogenetic and ontogenetic issues. In C. Lewis, & P. Mitchell (Eds.) *Children's early understanding of mind: Origins and development*, (pp. 19–45). New York: Cambridge University Press. Cited on page: 5, 134



- Moore, C., Bryant, D., & Furrow, D. (1989). Mental terms and the development of certainty. *Child Development*, 60(1), 167–171. Cited on page: 48
- Morgan, G., & Kegl, J. (2006). Nicaraguan sign language and theory of mind: The issue of critical periods and abilities. *Journal of Child Psychology and Psychiatry*, 47(8), 811–819. Cited on page: 34
- Moriguchi, Y., Okanda, M., & Itakura, S. (2008). Young children's yes bias: How does it relate to verbal ability, inhibitory control, and theory of mind? *First Language*, 28(4), 431–442. Cited on page: 111, 134
- Moses, L. J. (2001). Executive accounts of theory-of-mind development. *Child Development*, 72(3), 688–690. Cited on page: 12
- Naigles, L. R. (2000). Manipulating the input: Studies in mental verb acquisition. In B. Landau, J. Sabini, J. Jonides, & E. L. Newport (Eds.) *Perception, cognition, and language: essays in honor of Henry and Lila Gleitman*, (pp. 245–274). The MIT Press. Cited on page: 41
- Nelson, K. (2005). Language pathways into the community of minds. In J. W. Astington, & J. A. Baird (Eds.) *Why language matters for theory of mind*, (pp. 26–49). Oxford University Press, USA. Cited on page: 25, 26
- Neumann, A., Thoermer, C., & Sodian, B. (2008). *False belief understanding in 18-month-olds' anticipatory looking behavior: An eye-tracking study*. Talk given at XXIX International congress of psychology, Berlin, Germany. Cited on page: 18
- Newmeyer, F. J. (2010). What conversational english tells us about the nature of grammar: A critique of Thompson's analysis of object complements. *Language Usage and Language Structure*, (p. 3). Cited on page: 47
- Newton, A. M., & Villiers, J. G. (2007). Thinking while talking: Adults fail nonverbal false-belief reasoning. *Psychological Science*, 18(7), 574–579. Cited on page: 31
- Nixon, S. M. (2005). Mental state verb production and sentential complements in four-year-old children. *First Language*, 25(1), 19. Cited on page: 46, 47
- Noonan, M. (1985). Complementation. In T. Shopen (Ed.) *Language typology and syntactic description, vol. 2: Complex constructions*, (pp. 42–140). Cambridge: Cambridge University Press. Cited on page: 41, 44
- Norris, R., & Millan, S. (1991). *Theory of mind: New directions*. Social Psychology Seminar, University of Oxford. Cited on page: 14, 133
- Olson, D. R. (1988). On the origins of beliefs and other intentional states in children. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.) *Developing theories of mind*, (pp. 414–426). New York: Cambridge University Press. Cited on page: 26
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false

- beliefs? *Science*, *308*(5719), 255–258. Cited on page: 3, 14, 15, 20, 22
- Onishi, K. H., Baillargeon, R., & Leslie, A. M. (2007). 15-month-old infants detect violations in pretend scenarios. *Acta Psychologica*, *124*(1), 106–128. Cited on page: 21
- Papafragou, A. (2002). Mindreading and verbal communication. *Mind & Language*, *17*(1&2), 55–67. Cited on page: 2
- Papafragou, A., Cassidy, K., & Gleitman, L. (2007). When we think about thinking: The acquisition of belief verbs. *Cognition*, *105*(1), 125–165. Cited on page: 45
- Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of clinical epidemiology*, *49*(12), 1373–1379. Cited on page: 90
- Pérez-Leroux, A., & Schulz, P. (1999). The role of tense and aspect in the acquisition of factivity: children's interpretation of factive complements in English, German and Spanish. *First Language*, *19*(55), 29–54. Cited on page: 48
- Perner, J. (1991). *Understanding the representational mind*. MIT Press Cambridge, MA. Cited on page: 10, 12, 137
- Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, *5*(2), 125–137. Cited on page: 1, 11, 60, 65
- Perner, J., Rendl, B., & Garnham, A. (2007). Objects of desire, thought, and reality: problems of anchoring discourse referents in development. *Mind & Language*, *22*(5), 475–513. Cited on page: 10, 19, 21, 22, 97, 136
- Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: How deep? *Science*, *308*(5719), 214–216. Cited on page: 20, 22
- Perner, J., Sprung, M., Zauner, P., & Haider, H. (2003). Want that is understood well before say that, think that, and false belief: A test of de Villiers's linguistic determinism on German-speaking children. *Child Development*, *74*(1), 179–188. Cited on page: 3, 36, 46
- Perner, J., Stummer, S., Sprung, M., & Doherty, M. (2002). Theory of mind finds its piagetian perspective: Why alternative naming comes with understanding belief. *Cognitive Development*, *17*(3-4), 1451–1472. Cited on page: 37
- Perner, J., Zauner, P., & Sprung, M. (2005). What does "that" have to do with point of view? Conflicting desires and "want" in German. In J. W. Astington, & J. A. Baird (Eds.) *Why language matters for theory of mind*, (pp. 220–244). Oxford University Press, USA. Cited on page: 36, 37, 46
- Peterson, C., & Siegal, M. (1999). Representing inner worlds: Theory of mind in autistic, deaf, and normal hearing children. *Psychological Science*, *10*(2), 126–129. Cited on page: 34

- Peterson, C., Wellman, H., & Liu, D. (2005). Steps in theory-of-mind development for children with deafness or autism. *Child Development*, *76*(2), 502–517. Cited on page: 34
- Peterson, C. C., & Siegal, M. (2000). Insights into theory of mind from deafness and autism. *Mind & Language*, *15*(1), 123–145. Cited on page: 26, 34
- Phinney, M. (1981). Children's interpretation of negation in complex sentences. In S. L. Tavakolian (Ed.) *Language acquisition and linguistic theory*, (pp. 116–138). Cambridge, MA: MIT Press. Cited on page: 126
- Pinker, S., & Bloom, P. (1990). Natural selection and natural language. *Behavioral and Brain Sciences*, *13*(4), 707–784. Cited on page: 28
- Plaut, D. C., & Karmiloff-Smith, A. (1993). Representational development and theory-of-mind computations. *Behavioral and Brain Sciences*, *16*(1), 70–71. Cited on page: 13, 14, 133
- Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: suspiciously human? *Trends in Cognitive Sciences*, *7*(4), 157–160. Cited on page: 22
- Povinelli, D. J., & Vonk, J. (2004). We don't need a microscope to explore the chimpanzee's mind. *Mind & Language*, *19*(1), 1–28. Cited on page: 22
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 717–731. Cited on page: 100
- Premack, D. (2004). Is language the key to human intelligence? *Science*, *303*(5656), 318–320. Cited on page: 1
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain sciences*, *1*(4), 515–526. Cited on page: 9
- Pyers, J. E., & Senghas, A. (2009). Language promotes false-belief understanding. *Psychological Science*, *20*(7), 805–812. Cited on page: 34, 35
- Quine, W. V. (1956). Quantifiers and propositional attitudes. *The Journal of Philosophy*, *53*(5), 177–187. Cited on page: 27
- Quirk, R., Greenbaum, S., Leech, G., Svartvik, J., & Crystal, D. (1985). *A comprehensive grammar of the English language*. Longman London, UK. Cited on page: 43
- R Development Core Team (2009). R: A language and environment for statistical computing. URL <http://www.R-project.org> Cited on page: 73
- Ricard, M., Girouard, P. C., & Décarie, T. G. (1999). Personal pronouns and perspective taking in toddlers. *Journal of Child Language*, *26*(3), 681–697. Cited on page: 2
- Rosander, K., & von Hofsten, C. (2004). Infants' emerging ability to represent

- occluded object motion. *Cognition*, 91(1), 1–22. Cited on page: 16
- Rothweiler, M. (1993). *Der Erwerb von Nebensätzen im Deutschen: Eine Pilotstudie*. M. Niemeyer. Cited on page: 43, 46, 125
- Rozeboom, W. W. (1972). Problems in the psycho-philosophy of knowledge. In J. R. Royce, & W. W. Rozeboom (Eds.) *The psychology of knowing*, (pp. 25–55). New York: Gordon & Breach Publishing Group. Cited on page: 27
- Ruffman, T. (2000). Nonverbal theory of mind: Is it important, is it implicit, is it simulation, is it relevant to autism. In J. Astington (Ed.) *Minds in the making: Essays in honor of David R. Olson*, (pp. 250–266). Oxford: Blackwell. Cited on page: 16, 23
- Ruffman, T. (2004). *Theory of mind and language ability: Understanding the bigger picture*. Invited paper, web conference on coevolution of language and theory of mind. Retrieved from 06/21/2007, from <http://www.interdisciplines.org/coevolution/papers/8>. Cited on page: 3, 23, 35, 38, 136
- Ruffman, T., Garnham, W., Import, A., & Connolly, D. (2001a). Does eye gaze indicate implicit knowledge of false belief? Charting transitions in knowledge. *Journal of Experimental Child Psychology*, 80(3), 201–224. Cited on page: 3, 10, 16, 39, 53, 79, 96, 130, 132
- Ruffman, T., Garnham, W., & Rideout, P. (2001b). Social understanding in autism: Eye gaze as a measure of core insights. *Journal of Child Psychology and Psychiatry*, 42(8), 1083–1094. Cited on page: 32, 33
- Ruffman, T., & Perner, J. (2005). Do infants really understand false belief? Response to Leslie. *Trends in Cognitive Sciences*, 9(10), 462–463. Cited on page: 20, 22
- Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Development*, 73(3), 734–751. Cited on page: 3, 26
- Ruffman, T., Slade, L., Rowlandson, K., Rumsey, C., & Garnham, A. (2003). How language relates to belief, desire, and emotion understanding. *Cognitive Development*, 18(2), 139–158. Cited on page: 2, 3, 23, 30, 35, 54, 131
- Schick, B., de Villiers, P. A., de Villiers, J. G., & Hoffmeister, R. (2007). Language and theory of mind: A study of deaf children. *Child Development*, 78(2), 376–396. Cited on page: 2, 34
- Scholl, B. J., & Leslie, A. M. (2001). Minds, modules, and meta-analysis. *Child Development*, 72(3), 696–701. Cited on page: 10, 12, 137
- Schulz, P. (2002). The interaction of lexical-semantics, syntax and discourse in the acquisition of factivity. In B. Skarabela, S. Fish, & S. H.-J. Do (Eds.) *Proceedings of the 26th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press. Cited on page: 48

- Schulz, P. (2003). *Factivity: its nature and acquisition*. M. Niemeyer. Cited on page: 44, 48
- Schulz, P., & Ludwig, C. (2008). *The role of theory of mind in the acquisition of factivity*. Talk presented at the Symposium "Acquiring the semantics and syntax of presuppositions" XIth International Congress for the Study of Child Language (IASCL), Edinburgh, GB. Cited on page: 29, 48, 49, 50, 134
- Schulz, P., & Meissner, A. (2008). *Understanding Theory of Mind and complementation - The linguistic determinism hypothesis revisited*. Talk presented at the 28th Annual Conference on Language Development (BUCLD) Boston, USA. Cited on page: 48
- Scott, R. M., & Baillargeon, R. (2009). Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development*, 80(4), 1172–1196. Cited on page: 21, 22
- Sekerina, I., Stromswold, K., & Hestvik, A. (2004). How do adults and children process referentially ambiguous pronouns? *Journal of Child Language*, 31(1), 123–152. Cited on page: 111, 130
- Senghas, A. (1995). The development of Nicaraguan Sign Language via the language acquisition process. In D. MacLaughlin, & S. McEwen (Eds.) *Proceedings of the 19th Annual Boston University Conference on Language Development*, (pp. 543–552). Cited on page: 34
- Senghas, A. (2003). Intergenerational influence and ontogenetic development in the emergence of spatial grammar in Nicaraguan Sign Language. *Cognitive Development*, 18(4), 511–531. Cited on page: 34
- Senju, A., Southgate, V., Miura, Y., Matsui, T., Hasegawa, T., Tojo, Y., Osanai, H., & Csibra, G. (2010). Absence of spontaneous action anticipation by false belief attribution in children with autism spectrum disorder. *Development and Psychopathology*, 22(2), 353–360. Cited on page: 32
- Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes: An absence of spontaneous theory of mind in asperger syndrome. *Science*, 325(5942), 883–885. Cited on page: 32
- Shatz, M., Wellman, H., & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition*, 14(3), 301–321. Cited on page: 46, 47
- Siegal, M., & Peterson, C. C. (1996). Breaking the mold: A fresh look at questions about children's understanding of lies and mistakes. *Developmental Psychology*, 32(2), 322–34. Cited on page: 134
- Siegal, M., Varley, R., & Want, S. C. (2001). Mind over grammar: Reasoning in aphasia and development. *Trends in Cognitive Sciences*, 5(7), 296–301. Cited on page: 31

- Singer, M. S. (1987). A developmental study of children's comprehension of negation in complex sentences involving clauses. *The Journal of Genetic Psychology*, *118*(2), 189–196. Cited on page: 126
- Slade, L., & Ruffman, T. (2005). How language does (and does not) relate to theory of mind: A longitudinal study of syntax, semantics, working memory and false belief. *British Journal of Developmental Psychology*, *23*(1), 117–141. Cited on page: 36
- Slaughter, V., Peterson, C. C., & Mackintosh, E. (2007). Mind what mother says: Narrative input and theory of mind in typical children and those on the autism spectrum. *Child Development*, *78*(3), 839–858. Cited on page: 26
- Smith, M., Apperly, I., & White, V. (2003). False belief reasoning and the acquisition of relative clause sentences. *Child Development*, *74*(6), 1709–1719. Cited on page: 3, 30
- Sodian, B., & Thoermer, C. (2006). Theory of mind. In W. Schneider, & B. Sodian (Eds.) *Enzyklopädie der Psychologie, Themenbereich C, Serie V, Band 2: Kognitive Entwicklung*, (pp. 495–608). Göttingen: Hogrefe. Cited on page: 1
- Song, H., Onishi, K. H., Baillargeon, R., & Fisher, C. (2008). Can an agent's false belief be corrected by an appropriate communication? Psychological reasoning in 18-month-old infants. *Cognition*, *109*(3), 295–315. Cited on page: 21
- Southgate, V., Chevallier, C., & Csibra, G. (in press). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science*. Cited on page: 15
- Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*(7), 587–592. Cited on page: xi, 3, 10, 14, 15, 17, 18, 19, 53, 59, 79, 80, 97, 111, 130, 132
- Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, *18*(7), 580–586. Cited on page: 20, 128
- Swoboda, E. (2006). *Die Bedeutung spezifisch sprachlicher Erwerbsaspekte für die Entwicklung einer Theory of Mind. Eine empirische Studie zur First-Hypothesis*. Ph.D. thesis. Cited on page: 33
- Symons, D. K. (2004). Mental state discourse, theory of mind, and the internalization of self-other understanding. *Developmental Review*, *24*(2), 159–188. Cited on page: 26
- Tager-Flusberg, H. (1997). Language acquisition and theory of mind: Contributions from the study of autism. In L. B. Adamson, & M. A. Ronski (Eds.) *Research on communication and language disorders: Contributions to theories of language development*, (pp. 135–160). Baltimore, MD: Paul Brookes. Cited on page: 26
- Tager-Flusberg, H., & Joseph, R. M. (2005). How language facilitates the acquisition of false-belief understanding in children with autism. In J. W. Astington, &

- J. A. Baird (Eds.) *Why language matters for theory of mind*, (pp. 298–318). Oxford University Press, USA. Cited on page: 2, 32, 43, 55, 60, 70, 71, 99, 101, 131, 132, 134
- Tardif, T., So, C., & Kaciroti, N. (2007). Language and false belief: Evidence for general, not specific, effects in Cantonese-speaking preschoolers. *Developmental psychology*, 43(2), 318–340. Cited on page: 36
- Tardif, T., Wellman, H., & Cheung, K. (2004). False belief understanding in Cantonese-speaking children. *Journal of Child Language*, 31(4), 779–800. Cited on page: 3, 37
- Tardif, T., & Wellman, H. M. (2000). Acquisition of mental state language in Mandarin-and Cantonese-speaking children. *Developmental Psychology*, 36(1), 25–43. Cited on page: 3, 36
- Thompson, S. A. (2002). "Object complements" and conversation: Towards a realistic account. *Studies in Language*, 26(1), 125–163. Cited on page: 46, 47
- Thompson, S. A., & Mulac, A. (1991). A quantitative perspective on the grammaticalization of epistemic parentheticals in English. In E. C. Traugott, & B. Heine (Eds.) *Approaches to grammaticalization*, vol. 2, (pp. 313–339). Cited on page: 43
- Thornton, R., & Crain, S. (1994). Successful cyclic movement. In T. Hoekstra, & B. Schwartz (Eds.) *Language Acquisition Studies in Generative Grammar*, (pp. 215–253). Cited on page: 49, 51, 99
- Tomasello, M. J., M. and Farrar (1986). Joint attention and early language. *Child Development*, 57(6), 1454–1463. Cited on page: 2
- Träuble, B., Marinović, V., & Pauen, S. (2010). Early theory of mind competencies: Do infants understand others' beliefs? *Infancy*, 15(4), 434–444. Cited on page: 21
- Varley, R., & Siegal, M. (2000). Evidence for cognition without grammar from causal reasoning and theory of mind' in an agrammatic aphasic patient. *Current Biology*, 10(12), 723–726. Cited on page: 31
- Varley, R., Siegal, M., & Want, S. C. (2001). Severe impairment in grammar does not preclude theory of mind. *Neurocase*, 7(6), 489–493. Cited on page: 31
- Weissenborn, J., Höhle, B., Kiefer, D., & Cavar, D. (1998). Children's sensitivity to word-order violations in German: Evidence for very early parameter setting. In A. Greenhill, M. Hughes, H. Littlefield, & H. Walsh (Eds.) *22nd Annual Boston University Conference on Language Development, Vol. 2*, (pp. 756–767). Cited on page: 43, 125
- Wellman, H., & Bartsch, K. (1988). Young children's reasoning about beliefs. *Cognition*, 30(3), 239–277. Cited on page: 13
- Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.

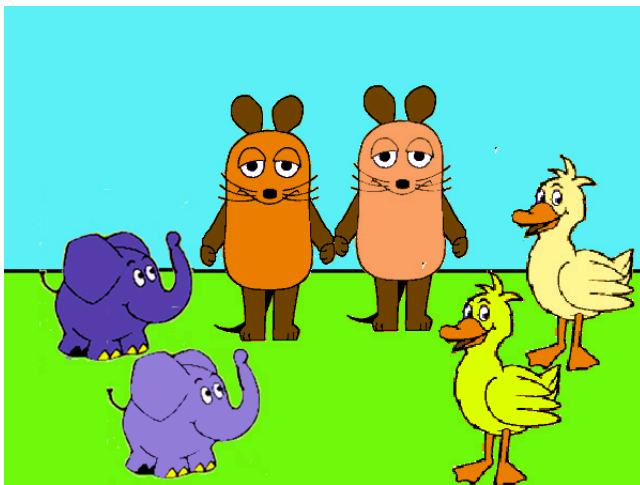
- Cited on page: 10, 12, 137
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*(3), 655–684. Cited on page: 10, 12, 15, 81, 97, 132
- Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, *75*(2), 523–541. Cited on page: 9, 127
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103. Cited on page: 1, 9, 11, 132
- Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*. Pears, D. F: London: Routledge, 1981. Cited on page: 1
- Wode, H. (1977). Four early stages in the development of LI negation. *Journal of Child Language*, *4*(1), 87–102. Cited on page: 124
- Woolfe, T., Want, S. C., & Siegal, M. (2002). Signposts to development: Theory of mind in deaf children. *Child Development*, *73*(3), 768–778. Cited on page: xi, 9, 34, 39
- Yi, L. (2009). *Why Do Young Children Fail in False Belief Tasks: Linguistic Representations and Implicit Processing*. Ph.D. thesis. Cited on page: 18, 19, 96, 97
- Zaitchik, D. (1991). Is only seeing really believing? Sources of the true belief in the false belief task. *Cognitive Development*, *6*(1), 91–103. Cited on page: 14, 133



# Appendix: Material

## A.1 Truth-Value-Judgment Task

Introductory picture

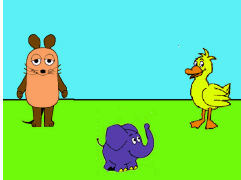
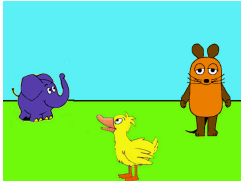
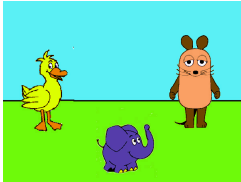
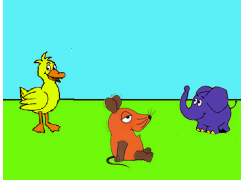


*Zeig mir mal die Enten!*  
'Please show me the ducks!'






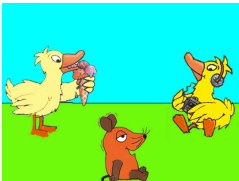
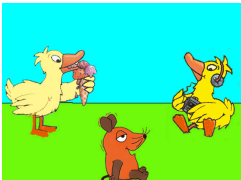
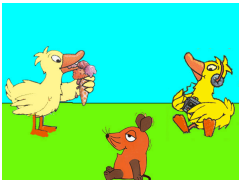
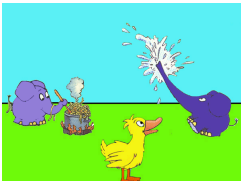
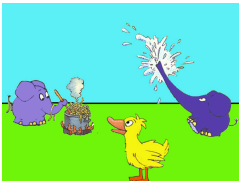
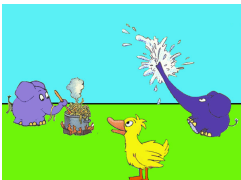
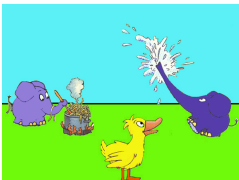
*Zeig mir mal die Mäuse!*  
'Please show me the mice!'

*Zeig mir mal die Elefanten!*  
'Please show me the elephants!'

## Practice items

No.	Sentence	Picture	Correct response
1	<i>Der Elefant sieht gerade die Ente.</i> 'The elephant is seeing the duck.'		yes
2	<i>Die Ente sieht gerade die Maus.</i> 'The duck is seeing the mouse.'		no
3	<i>Der Elefant sieht gerade die Ente.</i> 'The elephant is seeing the duck.'		no
4	<i>Die Maus sieht gerade den Elefanten.</i> 'The mouse is seeing the elephant.'		yes

## Test items

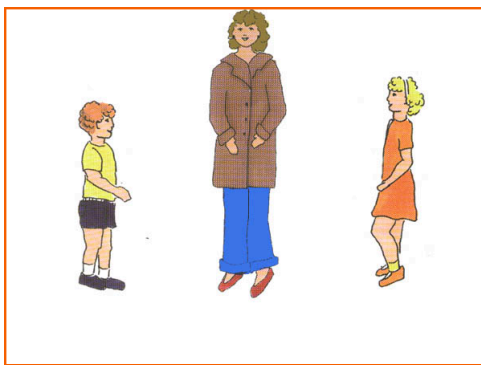
No.	Sentence	Picture	
		correct <i>yes</i> answer	correct <i>no</i> answer
1	<i>Der Elefant sieht gerade, dass die Maus Fußball spielt.</i>  'The elephant is seeing that the mouse is playing soccer.'		
2	<i>Der Elefant sieht gerade, dass die Maus Auto fährt.</i>  'The elephant is seeing that the mouse is driving a car.'		
3	<i>Die Maus sieht gerade, dass die Ente Eis isst.</i>  'The mouse is seeing that the mouse is eating ice cream.'		
4	<i>Die Maus sieht gerade, dass die Ente Musik hört.</i>  'The mouse is seeing that the duck is listening to music.'		
5	<i>Der Ente sieht gerade, dass der Elefant Wasser spritzt.</i>  'The duck is seeing that the elephant is splashing with water.'		
6	<i>Der Ente sieht gerade, dass der Elefant Nudeln kocht.</i>  'The duck is seeing that the elephant is cooking pasta.'		

## Order of test items

No.	Sentence	Correct answer	
		Version A	Version B
1	<i>Der Elefant sieht gerade, dass die Maus Fußball spielt.</i> 'The elephant is seeing that the mouse is playing soccer.'	yes	no
2	<i>Die Maus sieht gerade, dass die Ente Eis isst.</i> 'The mouse is seeing that the duck is eating ice cream.'	no	yes
3	<i>Die Ente sieht gerade, dass der Elefant Wasser spritzt.</i> 'The elephant is seeing that the mouse is splashing water.'	no	yes
4	<i>Der Elefant sieht gerade, dass die Maus Auto fährt.</i> 'The elephant is seeing that the mouse is driving a car.'	yes	no
5	<i>Die Maus sieht gerade, dass die Ente Musik hört.</i> 'The mouse is seeing that the duck is listening to music.'	no	yes
6	<i>Die Ente sieht gerade, dass der Elefant Nudeln kocht.</i> 'The duck is seeing that the elephant is cooking pasta.'	yes	no
7	<i>Der Elefant sieht gerade, dass die Maus Fußball spielt.</i> 'The elephant is seeing that the mouse is playing soccer.'	no	yes
8	<i>Die Maus sieht gerade, dass die Ente Eis isst.</i> 'The mouse is seeing that the duck is eating ice cream.'	yes	no
9	<i>Die Ente sieht gerade, dass der Elefant Wasser spritzt.</i> 'The elephant is seeing that the mouse is splashing water.'	yes	no
10	<i>Der Elefant sieht gerade, dass die Maus Auto fährt.</i> 'The elephant is seeing that the mouse is driving a car.'	no	yes
11	<i>Die Maus sieht gerade, dass die Ente Musik hört.</i> 'The mouse is seeing that the duck is listening to music.'	yes	no
12	<i>Die Ente sieht gerade, dass der Elefant Nudeln kocht.</i> 'The duck is seeing that the elephant is cooking pasta.'	no	yes

## A.2 Sentence-Picture Matching Task

### Introductory picture



*Guck mal, das sind Tim, Lisa und ihre Mama.*

*Ich zeige Dir Bilder, was sie so machen.*

*Aber es sind zu viele Bilder in den Computer gerutscht.*

*Nur ein Bild passt zur Geschichte.*

*Hilfst Du dem Erwin, das richtige Bild zu finden?*

'Look! This is Tim, Lisa and their mommy.'

'I will show some pictures to you what they are doing.'

'But there are to many pictures in the computer.'




'Just one picture is matching to a story.'


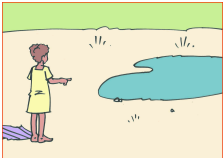


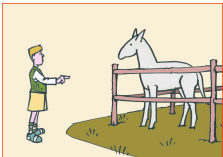
'Can you help Erwin to find the correct picture?'

## Test items

No.	Sentence	Picture
1	<p><i>Die Mama sieht, dass Lisa auf sie zeigt.</i></p> <p>‘Mommy sees that Lisa is pointing at her.’</p>	
2	<p><i>Tim sieht, dass die Mama sich sieht.</i></p> <p>‘Tim sees that Mommy is seeing herself.’</p>	
3	<p><i>Die Mama sieht, dass Lisa auf sich zeigt.</i></p> <p>‘Mommy sees that Lisa is pointing at herself.’</p>	
4	<p><i>Tim sieht, dass die Mama sie sieht.</i></p> <p>‘Tim sees that Mommy is seeing her.’</p>	

## A.3 Reporting-the-Complement Task

No.	Type	Sentences	Picture
Practice Items			
1	TRUE	<p><i>Das ist Peter.</i>  <i>Peter sagte, dass die Sonne scheint. (C)</i>  <i>Draußen war wirklich schönes Wetter. (S)</i>            ‘This is Peter.’            ‘Peter said that the sun is shining.’ (C)            ‘The weather was really nice outside.’ (S)</p>	
2	FALSE	<p><i>Das ist Julia.</i>  <i>Julia sagte, dass die Brücke kaputt ist. (C)</i>  <i>Die Brücke war in Wirklichkeit ganz. (S)</i>            ‘This is Julia.’            ‘Julia said that the bridge is broken.’ (C)            ‘The bridge was really entire.’ (S)</p>	
Test Items			
3	TRUE	<p><i>Das ist Marie.</i>  <i>Marie sagte, dass sie eine Banane möchte. (C)</i>  <i>Sie hatte wirklich Hunger auf Obst. (S)</i>            ‘This is Marie.’            ‘Marie said that she wants to have a banana.’ (C)            ‘She was really hungry for fruits.’ (S)</p>	

No.	Type	Sentences	Picture
4	TRUE	<p><i>Das ist Tommi.</i>  <i>Tommi sagte, dass er friert. (C)</i>  <i>Im Haus war es wirklich kalt. (S)</i>            ‘This is Tommi.’            ‘Tommi said that he is freezing.’ (C)            ‘It was really cold in the house.’ (S)</p>	
5	FALSE	<p><i>Das ist Anna.</i>  <i>Anna sagte, dass das Wasser tief ist. (C)</i>  <i>Das Wasser war in Wirklichkeit flach. (S)</i>            ‘This is Anna.’            ‘Anna said that the water is deep.’ (C)            ‘The water was really flat.’ (S)</p>	
6	FALSE	<p><i>Das ist Benny.</i>  <i>Benny sagte, dass im Sand ein Stift liegt. (C)</i>  <i>Im Sand lag in Wirklichkeit ein Löffel. (S)</i>            ‘This is Benny.’            ‘Benny said that there is a pencil in the sand.’ (C)            ‘There was really a spoon in the sand.’ (S)</p>	
7	TRUE	<p><i>Das ist Steffi.</i>  <i>Steffi sagte, dass das Bonbon süß schmeckt. (C)</i>  <i>Das Bonbon war wirklich lecker. (S)</i>            ‘This is Steffi.’            ‘Steffi said that the candy is tasting sweet.’ (C)            ‘The candy was really tasty.’ (S)</p>	
8	FALSE	<p><i>Das ist Tobi.</i>  <i>Tobi sagte, dass er einen Esel sieht. (C)</i>  <i>Tobi sah in Wirklichkeit ein Pferd. (S)</i>            ‘This is Tobi.’            ‘Tobi said that he is seeing a donkey.’ (C)            ‘Tobi saw really a horse.’ (S)</p>	



## A.4 Comprehension of Negation: Sentence-Picture-Matching Task

### Practice items

No.	Item	Picture
1	<i>Schuh</i> 'Shoe'	
2	<i>Vogel</i> 'Bird'	
3	<i>Kamm</i> 'Comb'	
4	<i>Apfel</i> 'Apple'	

## Test items

No.	Item	Picture
1	<i>Das Mädchen springt nicht.</i> 'The does not jump.'	
2	<i>Der Hund sitzt nicht.</i> 'The dog does not sit.'	
3	<i>Der Junge läuft nicht.</i> 'The boy does not run.'	
4	<i>Der Hund trinkt nicht.</i> 'The dog does not drink.'	

# Erklärung

Ich versichere, diese Arbeit selbstständig und ohne unzulässige Hilfe Dritter verfasst zu haben. Bei der Abfassung der Dissertation habe ich nur die angegebenen Hilfsmittel benutzt. Alle inhaltlich oder wörtlich übernommenen Aussagen habe ich als solche gekennzeichnet.

Ich versichere weiterhin, dass ich diese Dissertation weder an der gegenwärtigen noch in einer anderen Fassung einer anderen Fakultät einer wissenschaftlichen Hochschule zur Begutachtung im Rahmen eines Promotionsverfahrens vorgelegt habe.

Potsdam, den 08.11.2010

