

University of Groningen

Between direct and indirect speech

Köder, Franziska Maria

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2016

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Köder, F. M. (2016). Between direct and indirect speech: The acquisition of pronouns in reported speech
Groningen: University of Groningen

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Between direct and indirect speech



Groningen Dissertations in Linguistics 149
© Franziska Köder, 2016

ISSN: 0928-0030

ISBN printed version: 978-90-367-8477-1

ISBN digital version: 978-90-367-8476-4

The work presented in this thesis has been carried out under the auspices of the Faculty of Philosophy and the Center for Language and Cognition Groningen (CLGC) of the University of Groningen.

This research was supported by the EU under FP7, ERC Starting Grant 263890-BLENDS (Emar Maier).

Cover design by Maarten Bruinsma and Franziska Köder
Typesetting by *studio Michał Sławiński, thesisprint.eu
Printed in Poland



university of
 groningen

Between direct and indirect speech

The acquisition of pronouns in reported speech

PhD thesis

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus Prof. E. Sterken
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Thursday 18 February 2016 at 16.15 hours

by

Franziska Maria Köder

born on 3 November 1985
in Schwäbisch Hall, Germany

Supervisors

Prof. J.W. Romeijn

Prof. P. Hendriks

Co-supervisor

Dr. E. Maier

Assessment Committee

Prof. E. Engberg-Pedersen

Prof. M. Steinbach

Prof. L.C. Verbrugge

ACKNOWLEDGMENTS

Coming to Groningen has been one of the best decisions of my life. I have not only learned a lot about linguistics and philosophy, but I have also met many wonderful people. I want to take the opportunity to thank some of them here.

I was very fortunate to get offered a PhD position in Emar Maier's ERC project BLENDS. First of all, I would like to express my gratitude to Emar, my primary supervisor and co-promotor. Emar, you have helped me tremendously in all stages of my PhD. You gave me the freedom to pursue my own interests, while providing me with enough guidance not to lose my way. Over the years, you have commented on countless drafts of articles, chapters and presentations, and you have introduced me to the customs and practices of academia. I have learned so much from you that it is hard to imagine what my thesis would look like without you. Thanks for everything!

I am also very grateful for having Petra Hendriks as my promotor. Petra, your detailed comments on chapters and articles have been invaluable. I loved our discussions on pronouns and experimental testing, and I am especially grateful that you offered me the opportunity to teach some classes on semantics and irony. You are right that pronouns are addictive and I will certainly come back to *you* at some point.

I am thankful that Jan-Willem Romeijn agreed to be my promotor in the final year of my PhD.

Many people have helped me planning, preparing and executing the three experiments that I report in this thesis. I would like to thank Jelmer van der Linde for programming and Martijn Luinstra for graphically designing experiments 1 and 3; Han Thomas Adriaansen, Bart Hollebrandse and Emar Maier for lending their voices to Monkey, Elephant and Dog, respectively; Iris Strangmann and Anna de Koster for assisting with testing children and adults; Martijn Wieling for help with the statistical analysis; the Westerschool (Wildervank) – especially Ruby Trebels –, OBS De Sleutel (Wildervank) and OBS De Sterrensteen (Groningen) for their cooperation; and crucially, the many children and adults for participating in the experiments.

I feel privileged that I could discuss my ideas with both linguists and philosophers. I thank the members of the *Acquisition Lab* and the *Semantics & Cognition* group for discussing my experimental designs and results with me, especially Burcu Arslan, Gisi Cannizzaro, Petra Hendriks, John Hoeks, Bart Hollebrandse, Sanne Kuijper, Jessica Overweg, Atty Schouwenaars, Jennifer Spenader, Laurie Stowe, Angeliek van Hout, Jacolien van Rij and Margreet Vogelzang. At the Faculty of Philosophy, I thank the varying members of the PCCP and the WiP meetings for interesting talks and discussions, and the “lunch regulars”, among others, Allard, Annet, Barteld, Bouke, Coos, Emar, Fré, Han Thomas, Harrie, Herman, Job, Leah, Marc, Marta, Nynke, Sander, and Sanne, for the *gezelligheid*. In both faculties, I had great office mates with Carlos, Chantelle, Dave, Gisi, Jacolien, Jessica, Marc, Margreet, Marijana, Paolo, Paulien, Sanne and Stijn.

I am grateful for the opportunity to attend many interesting conferences in different countries and to share the local pleasures with, for instance, Karolina (Saint Petersburg), Margreet (Trondheim, Edinburgh) and Ditte (Leipzig). I am indebted to the audience of these conferences, the anonymous referees of the articles some chapters of this dissertation are based on, and the members of the reading committee for many helpful questions, comments, and suggestions. One of the highlights of my time as a PhD student was organizing the conference *(Re)presenting the Speech of Others* together with Emar. Thanks to all participants for the inspiring talks and discussions!

During my time in Groningen, I have met many interesting and kind people, but I would like to single out a few. A warm thanks goes to the Maier/Spenader family for making me feel welcome in Groningen. Jennifer, you were always there for me with help and advice, and I am grateful that you offered me the opportunity to co-supervise student research projects on reported speech in Frisian-Dutch bilingual speakers and elderly people. Kaylee and Kiki, thanks so much for pre-testing all my experiments and for making me beautiful drawings. Paulien, Marijana and Karolina, I am glad that you have shared the joys, doubts and frustrations of my PhD life with me and that you have distracted me from work with drinks, dinners, sleepovers, Bootcamp and climbing. *Tige tank*, Bruinsma family, for your warmth and incredible hospitality.

Finally, I would like to thank my parents, my sister Christine and my brother Philipp for their unconditional love and support. Even if we are often far apart, you are still so close to me. Maarten, I am so grateful that you are always there for me. Life with you is – sometimes literally – an adventure and I cannot wait to find out what's coming next.

LIST OF PUBLICATIONS

Parts of chapters 2, 3 and 4 have appeared previously in:

Köder, F. (2013). How children acquire reported speech in German and Dutch: A corpus study. In B. Sonnenhauser, C. Trautmann, & P. Noel (Eds.), „*Perspektiven*“. *Diskussionsforum Linguistik in Bayern / Bavarian Working Papers in Linguistics* 2 (pp. 15–28). München, Bamberg.

Köder, F., Maier, E., & Hendriks, P. (2015). Perspective shift increases processing effort of pronouns: A comparison between direct and indirect speech. *Language, Cognition and Neuroscience*, 30(8), 940–946. doi:10.1080/23273798.2015.1047460

Köder, F., & Maier, E. (2015). Children mix direct and indirect speech: Evidence from pronoun comprehension. *Journal of Child Language*. doi:10.1017/S0305000915000318

TABLE OF CONTENTS

1 INTRODUCTION	15
2 THEORETICAL BACKGROUND	17
2.1 SPEECH REPORTS	17
2.2 THE DIRECT–INDIRECT DISTINCTION	18
2.2.1 Direct and indirect speech in English, Dutch and German	18
2.2.2 Syntactic restrictions in indirect speech	20
2.2.3 Opacity vs. transparency: <i>wh</i> -movement and NPI licensing	21
2.2.4 Interpretation of context-dependent expressions	22
2.2.5 Direct speech: A verbatim rendition?	23
2.2.6 Prosodic differences	24
2.2.7 The vividness of direct speech	25
2.2.8 Different discourse functions	26
2.3 BETWEEN DIRECT AND INDIRECT SPEECH	27
2.3.1 V2 constructions	28
2.3.2 Mixed quotation	28
2.3.3 Free indirect discourse	29
2.3.4 Perspective-dependent expressions	29
2.3.5 Role shift in sign languages	30
2.3.6 Person magnetism	31
2.3.7 Towards a more fluid picture	31
2.4 CHILDREN’S ACQUISITION OF DIRECT AND INDIRECT SPEECH	33
2.4.1 General overview of children’s acquisition of speech reports	33
2.4.2 Acquisition of speech reports in Dutch and German	34
2.4.3 Marking of perspective shifts	36
2.4.4 Direct–indirect mixes	36
2.4.5 Children’s comprehension of complement clauses	37
2.4.6 <i>Wh</i> -movement	39
2.4.7 Prosodic quotation marking	40
2.4.8 Written quotation marking	41
2.4.9 Distinguishing different discourse functions	42
2.4.10 Children’s difficulty with the direct–indirect distinction	44
2.5 PRONOUNS IN SPEECH REPORTS	44
2.5.1 Person distinctions	45
2.5.2 Deictic vs. anaphoric pronouns	46
2.5.3 Quoted pronouns	47
2.5.4 Pronouns in Dutch and German	47

2.6 CHILDREN'S ACQUISITION OF DEICTIC PRONOUNS	48
2.6.1 Prerequisites of pronoun acquisition	48
2.6.2 Children's production and comprehension of pronouns	50
2.6.3 Acquisition of pronominal gender marking	52
2.6.4 Challenges of pronoun interpretation in speech reports	52
2.7 RESEARCH QUESTIONS AND OVERVIEW OF EMPIRICAL STUDIES	53
3 CORPUS STUDY: CHILDREN'S SPONTANEOUS PRODUCTION OF SPEECH	
REPORTS	57
3.1 DEVELOPMENTAL TRAJECTORY AND FREQUENCY	57
3.1.1 Introduction	57
3.1.2 Data and Methods	58
3.1.3 Children's first speech reports	61
3.1.4 Frequency of direct and indirect speech	64
3.1.4.1 General overview	64
3.1.4.2 Direct speech dominates	67
3.1.4.3 Indirect speech with verb-second and verb-final word order	69
3.1.4.4 Not many direct-indirect mixes	70
3.1.5 Input analysis: Reports of children's interaction partners	72
3.2 ZOOMING IN ON THE PRONOUN CORPUS: THREE PILOT STUDIES	75
3.2.1 The pronoun corpus	75
3.2.2 Study 1: Children's use of pronouns in speech reports	76
3.2.3 Study 2: Whose speech children report	78
3.2.4 Study 3: The influence of the discourse context	79
3.3 SUMMARY OF SPONTANEOUS SPEECH ANALYSIS	83
4 EXPERIMENT 1: WHO GETS THE BALL? A TABLET GAME TO TEST REPORTED	
SPEECH COMPREHENSION	85
4.1 ADULTS' INTERPRETATION OF PRONOUNS IN SPEECH REPORTS	85
4.1.1 Introduction and hypotheses	85
4.1.2 Method	87
4.1.2.1 Participants	87
4.1.2.2 Stimuli and Procedure	87
4.1.3 Results	91
4.1.3.1 Accuracy	91
4.1.3.2 Decision times	93
4.1.4 Discussion	95

4.2 CHILDREN'S INTERPRETATION OF PRONOUNS IN SPEECH REPORTS	97
4.2.1 Introduction	97
4.2.2 Hypotheses	98
4.2.3 Method	99
4.2.3.1 Participants	99
4.2.3.2 Stimuli and Procedure	100
4.2.4 Results	100
4.2.4.1 'No report' condition: Accuracy	100
4.2.4.2 'No report' condition: Decision times	101
4.2.4.3 'Speech report' condition: Accuracy	102
4.2.4.4 'Speech report' condition: Decision times	104
4.2.4.5 Direct-indirect mixes	104
4.2.5 Discussion	104
4.2.6 Conclusion	109
5 EXPERIMENT 2: TESTING THE COMPREHENSION AND PRODUCTION OF REPORTED SPEECH IN INTERACTION	111
5.1 INTRODUCTION AND HYPOTHESES	111
5.2 METHODS	113
5.2.1 Participants	113
5.2.2 Stimuli and Procedure	113
5.3 RESULTS	116
5.3.1 Comprehension	117
5.3.1.1 Accuracy of pronoun interpretation	117
5.3.1.2 Direct-indirect mixes	119
5.3.1.3 Comparison with experiment 1	119
5.3.2 Production	122
5.3.2.1 Overview of produced speech reports	122
5.3.2.2 Direct-indirect mixes	124
5.3.2.3 Addressee-oriented reports	125
5.4 DISCUSSION	127
5.4.1 Comprehension	127
5.4.2 Excursus: Special properties of 'you'	128
5.4.3 Production	131
5.4.4 Relation between comprehension and production	133
5.5 CONCLUSION	134

6 EXPERIMENT 3: TESTING THE COMPREHENSION OF REPORTED SPEECH IN A NARRATIVE CONTEXT	137
6.1 INTRODUCTION AND HYPOTHESES	137
6.2 METHOD	139
6.2.1 Participants	139
6.2.2 Stimuli and Procedure	139
6.3 RESULTS	144
6.3.1 Gender pre-test	144
6.3.2 Speech report test: Accuracy results	145
6.3.3 Speech report test: Comparison with experiment 1	146
6.3.4 Memory test	148
6.4 DISCUSSION	149
6.5 CONCLUSION	152
7 CONCLUSION	155
7.1 INTRODUCTION	155
7.2 FACTORS INFLUENCING PRONOUN INTERPRETATION IN SPEECH REPORTS	156
7.2.1 Discourse context	156
7.2.2 Role in the interaction	158
7.2.3 Anaphora and linguistic mention	159
7.3 A PRODUCTION–COMPREHENSION ASYMMETRY?	160
7.4 METHODOLOGICAL REMARKS	161
7.5 FUTURE RESEARCH	162
APPENDICES	167
APPENDIX A: EXPERIMENT 1	167
APPENDIX B: EXPERIMENT 2	169
APPENDIX C: EXPERIMENT 3	171
REFERENCES	181
SAMENVATTING	197
ABOUT THE AUTHOR	203
GRONINGEN DISSERTATIONS IN LINGUISTICS (GRODIL)	205

1

INTRODUCTION

Imagine that a mother tells her daughter Lucy on Thursday during dinner that she is not allowed to swim in the river. When Lucy's friend Robin asks her the next day at school to go swim in the river later, Lucy reports her mother's utterance:

- (1) My mom said that I am not allowed to swim in the river.

Speech reports like (1) are an essential part of human communication. We can encounter them almost everywhere: in everyday conversations, news reports, philosophical treatises and works of fiction. It is therefore not surprising that speech reports have attracted the interest of researchers outside the discipline of linguistics. Philosophers have, for instance, pondered about the meaning of words inside of quotation marks (e.g., Davidson, 1979; Frege, 1892). Literary scholars have studied the representation of speech and thought in literary texts (e.g., Fludernik, 1993). Psychologists have focused on how reported speech is mentally represented (Eerland, Engelen, & Zwaan, 2013; Yao & Scheepers, 2011).

Traditionally, two different types of speech reports are distinguished: direct speech (*oratio recta*) and indirect speech (*oratio obliqua*). In the situation described above, Lucy used indirect speech – repeated below for convenience. But she could just as well have reported her mother's utterance with a direct speech report such as (2). This would, for instance, allow her to mock her mother's tone of voice.

- (1) Indirect speech: My mom said that I am not allowed to swim in the river.
 (2) Direct speech: My mom said, "You're not allowed to swim in the river".

If we contrast the direct and indirect version of the same reported utterance, one crucial semantic difference becomes apparent: Lucy uses different pronouns to refer to herself: the first-person pronoun *I* in the indirect speech report (1), and the second-person pronoun *you* in the direct speech report (2). This difference in pronoun use is related to the perspective from which Lucy presents her mother's previous utterance. In indirect speech, Lucy reports the utterance from her own current perspective and therefore refers to herself with a first-person

pronoun. In direct speech, Lucy shifts to the perspective that her mother occupied when originally producing the utterance *You're not allowed to swim in the river*. From her mother's point of view in this earlier conversation, Lucy is the addressee and hence referent of *you*.

If the listener Robin is to correctly interpret pronouns like *I* and *you* in a speech report, she needs to know from which perspective the reported utterance is presented, i.e., she must be able to distinguish direct from indirect speech. Otherwise, Robin might for instance falsely assume that the second-person pronoun *you* in the direct speech report (2) refers to her, the current addressee instead of to Lucy, the original addressee. Previous developmental studies indicate that children have difficulties to clearly distinguish between direct and indirect speech (Goodell & Sachs, 1992; Hickmann, 1993; Hollebrandse, 2007). But it is unclear whether this also affects children's comprehension of pronouns in direct and indirect speech.

Despite the fact that the interpretation of pronouns and other indexicals in speech reports is a much discussed topic in semantics (see Schlenker, 2011, for an overview), we know little about the cognitive processes underlying the interpretation of context-dependent expressions in direct and indirect speech. Is it easier for a listener to interpret pronouns in direct or indirect speech? Do pragmatic factors such as the communicative situation influence the processing of speech reports? When do children learn to correctly interpret and produce pronouns in direct and indirect speech? I will answer these fundamental questions in this dissertation, based on a corpus study (chapter 3) and three psycholinguistic experiments (chapters 4, 5 and 6). More specifically, I look at the interpretation and production of speech reports in Dutch children and adults. With my research on the processing of speech reports, I hope to contribute to the empirical foundations of theories on reported speech in various disciplines such as linguistics, philosophy, literary studies and psychology.

Before I present my empirical studies, I will first provide the theoretical background against which these studies have to be understood. I start with a definition of speech reports (2.1). Next, I discuss whether it is justified to construe direct and indirect speech as two fundamentally different report types (2.2). Based on many examples of speech reports that contain features of both direct and indirect speech, I argue for a less rigid direct–indirect distinction (2.3). Subsequently, I present previous research on children's acquisition of direct and indirect speech (2.4). Next, I turn to pronouns and discuss relevant distinctions in the pronominal system (2.5). Afterwards, I summarize previous research on children's production and comprehension of deictic pronouns (2.6). The theoretical background chapter closes with a detailed overview of the empirical studies (2.7).

2

THEORETICAL BACKGROUND¹

2.1 SPEECH REPORTS

Speech reports come in a wide variety of forms and functions. Crucially, not all speech reports refer to a unique speech event in the past. Speakers can, for instance, also ascribe hypothetical utterances to others (see (3)), they can report what has not been said (see (4)) and summarize a series of utterances using placeholders like *such-and-such* (see (5)) (Koev, 2015; Sams, 2010; Tannen, 2007; Von Roncador, 1988).

- (3) Dad would probably say, “You’re old enough to swim in the river”.
- (4) Mom never said that I am not allowed to wade through the river.
- (5) Mom always says, “You’re not allowed to do such-and-such”.
- (6) Mom’s look was like, “Don’t you dare!”.

Speech can even be ascribed to entities that are clearly unable to communicate in a human language, such as animals or inanimate objects. Report (6) is an example of such a “fictive interaction” (Pascual, 2014): The reporting speaker pretends that the look on her mother’s face is saying *Don’t you dare!*.

In this dissertation, I use the term *speech report* in a wide sense for all kinds of explicit speech ascriptions, without presupposing the existence of an original utterance. I prefer *speech report* to other terms such as “constructed dialogue” (Tannen, 2007) or “reporting discourse” (Sakita, 2002) because it is more common and indicates that I confine myself to attributions of speech, leaving aside the closely related phenomena of reported thought, perception and action (Janssen & Van der Wurff, 1996; Lillo-Martin, 2012).

¹ Parts of chapter 2 are based on Köder and Maier (2015) and Köder (2013).

2.2 THE DIRECT–INDIRECT DISTINCTION

Speech reports are traditionally divided into two fundamentally different types: direct speech and indirect speech. Linguists typically consider direct speech a form of quotation (i.e., a form of reference to linguistic objects like sentences or utterances) and indirect speech an intensional clausal embedding, syntactically and semantically on a par with attitude ascriptions (*believes that*) and modal operators (*it is possible that*) (Kaplan, 1989; Maier, 2009; Zimmermann, 1991). According to the influential demonstration theory of quotation, the difference between direct and indirect speech is due to the fundamental opposition between demonstration and description: Direct quotations are a type of demonstration that depict certain aspects of an utterance (such as its content, the accent or emotional state of the speaker) and give the audience an impression what it would be like to listen to the reported speaker directly. Indirect speech reports, however, *describe* aspects of the reported utterance, focusing mainly on its propositional content (Clark & Gerrig, 1990; Davidson, to appear; Recanati, 2001; Wade & Clark, 1993).

But are direct and indirect speech really two clearly distinct types to report speech? In the following sections, I will search for criteria that distinguish direct and indirect speech on different levels of linguistic analysis, such as syntax, semantics, prosody and pragmatics.

2.2.1 Direct and indirect speech in English, Dutch and German

Syntactically, the main difference between direct and indirect speech is the level of syntactic integration. In indirect speech, the reported utterance is syntactically integrated in the linguistic environment, for instance in the form of a subordinate clause. By contrast, direct quotations of assertions are relatively independent and behave similarly to main clauses (Banfield, 1973; Li, 1986; Oshima, 2006).

Let us have a closer look at the syntactic differences between direct and indirect speech in English, Dutch and German. In English, the reported utterance *I am happy* retains verb-second word order in both direct speech (see (7a)) and indirect speech (see (7b)).

- (7) a. Anna said, “I am happy”.
 b. Anna said (that) she was happy.

The complementizer *that* in English is optional in indirect speech. Moreover, English exhibits sequence of tense in indirect speech. This means that the tense of the complement clause is adjusted to agree with the tense of the verb in the matrix clause. We can see this in (7b), where the original utterance *I am happy* (present tense) is changed to *that she was happy* (simple past) because the matrix clause is past tense (*said*).

In contrast to English, the direct–indirect distinction is syntactically more clearly marked in Dutch.

- (8) a. Anna zei: “Ik ben blij”.
 Anna say-PST I be-PRES happy
 ‘Anna said: “I am happy”’
 b. Anna zei dat ze blij was.
 Anna say-PST that she happy be-PST
 ‘Anna said that she was happy.’

In direct speech, assertions like *Ik ben blij* ‘I am happy’ have verb-second word order, as typical for Dutch main clauses. In indirect speech, however, the reported utterance has the form of a subordinate clause with verb-final word order and an obligatory complementizer *dat* ‘that’. Indirect speech reports in Dutch can also exhibit sequence of tense (Boogart, 1996).

In German, word order and complementizer do not always differentiate between direct and indirect speech. As in Dutch, direct reports of assertions have verb-second word order in German (see (9a)). Moreover, German also has a form of indirect speech with verb-final word order and complementizer (see (9b)).

- (9) a. Anna sagte: „Ich bin glücklich“.
 Anna say-PST I be-PRES happy
 ‘Anna said: “I am happy”’
 b. Anna sagte, dass sie glücklich ist.
 Anna say-PST that she happy be-PRES
 ‘Anna said that she is happy.’
 c. Anna sagte, sie sei glücklich.
 Anna say-PST she be-SBJV happy
 ‘Anna said she is happy.’

But unlike standard Dutch, German offers an additional indirect speech construction (see (9c)), that is sometimes called “abhängiger Hauptsatz” ‘dependent main clause’ (Auer, 1998) or “unselbständiger Verbzweit-Satz” ‘non-independent verb-second sentence’ (Reis, 1997). These terms reveal that this indirect verb-second construction² is situated between coordination and subordination. On the one hand, the report complement is dependent on the reporting clause. On the other hand, it is syntactically less integrated than the indirect verb-final construction because it lacks a complementizer and exhibits verb-second word order, which

² For a discussion why this reporting construction is a form of indirect speech rather than direct speech, see Köder (2013). There I show that dependent main clause constructions exhibit certain main clause phenomena such as the possible topicalization of various constituents and subject-auxiliary inversion in questions. However, semantically they behave as expected from indirect reports: deictic expressions have to be evaluated with respect to the reporting context, and they allow grammatical dependencies between reporting clause and report (e.g., the licensing of negative polarity items).

is typical for main clauses in German. In order to still mark the indirectness of the report, indirect verb-second constructions often contain verbs in the subjunctive mood (Fabricius-Hansen, 2002; Plank, 1986).³

The syntactic analysis shows that of these three languages, Dutch is the one that marks the direct–indirect distinction most clearly in the grammar. In Dutch, direct and indirect speech can be distinguished with respect to word order, sequence of tense and the presence/absence of a complementizer. Since Dutch leaves very little room for ambiguous reports, it is ideally suited for testing children’s and adults’ comprehension of pronouns in direct and indirect speech. This is why I chose Dutch as language of investigation for the experimental studies (discussed in chapters 4, 5 and 6). In the corpus study (chapter 3), I look at both Dutch and German children’s spontaneous use of speech reports. German is interesting because of its additional indirect verb-second construction, which resembles direct reports of assertions. In the following sections, I illustrate differences between direct and indirect speech mainly with examples from Dutch.

2.2.2 Syntactic restrictions in indirect speech

The fact that the reported utterance is syntactically integrated in indirect speech entails that the reported utterance is subject to syntactic and lexical restrictions typical for subordinate clauses. This explains why it is not possible to shift to English in a Dutch indirect speech report like (10a).

- (10) a. *Anna zei dat she is happy.
 ‘Anna said that she is happy.’
 b. Anna zei: “I am happy”.
 ‘Anna said: “I am happy”.’

In direct speech, however, the reported utterance within quotation marks is basically unconstrained. This means that not only utterances in a different language or dialect can be quoted (see (10b)), but also expressives like *ouch* (e.g., *Anna said: “Ouch!”*). According to Coulmas, “expressive elements are one feature which allows us to recognize direct speech” (Coulmas, 1985, p. 44).

The greater syntactic and lexical restrictions of indirect speech also come into view when reporting questions and requests. Compare, for instance, the direct and the indirect report of Anna’s command *Ga naar huis, Jan* ‘Go home, Jan’.

- (11) a. Anna zei: “Ga naar huis, Jan!”
 ‘Anna said: “Go home, Jan!”’

³ Indirect verb-final constructions in German can also contain verbs in the subjunctive. The subjunctive can signal that the reporting speaker distances himself epistemically from the reported utterance, withholding responsibility for the truth of the expressed proposition (Fabricius-Hansen, 2002; Plank, 1986).

- b. Anna zei tegen Jan dat hij naar huis moest gaan.
 ‘Anna said to Jan that he had to go home.’

In the direct speech report (11a), the original utterance including the imperative and vocative can be quoted without further adjustments. But in the indirect report (11b), the illocutionary force of the command is expressed with the modal auxiliary verb *moeten* ‘have to’. Vocatives cannot be reported at all in an indirect speech report (Banfield, 1973).

Taken together, shifts of language or dialect, certain kinds of expressives (e.g., *ouch*), imperatives and vocatives are usually cues for direct speech.

2.2.3 Opacity vs. transparency: *wh*-movement and NPI licensing

Another difference between direct and indirect speech is that indirect speech is syntactically and semantically transparent to grammatical processes, while direct speech is opaque. This is the reason why only indirect speech allows certain grammatical dependencies between reporting clause and report.

First, while direct speech blocks *wh*-extraction out of the quotation (12b), indirect speech allows extraction out of the complement (12c) (Schlenker, 2011):

- (12) a. Original: Jullie moeten je kamer opruimen.
 ‘You guys have to tidy up your room.’
 b. Direct: (*)Wat zei mama, “Jullie moeten doen”?
 ‘What did mom say, “You guys have to do”?’
 c. Indirect: Wat zei mama dat we moeten doen?
 ‘What did mom say (that) we have to do?’

Similarly, only in indirect speech does a main clause negation license negative polarity items (NPI) in the complement (Anand & Nevins, 2004).

- (13) a. Anna heeft nooit gezegd dat ook maar iemand ontslagen zou worden.
 ‘Anna has never said that anyone would get fired.’
 b. (*)Anna heeft nooit gezegd : “Ook maar iemand zal ontslagen worden”.
 ‘Anna did not say: “Anyone will get fired”’

In example (13), the Dutch NPI *ook maar* ‘any’ is licensed by the negation *nooit* ‘never’ in the indirect speech report (13a), but not in the direct speech report (13b). Inside of a quotation, NPIs are not in the scope of the matrix clause negation.

Wh-extraction and the licensing of NPIs provide us with a syntactic and a semantic test to distinguish direct from indirect speech.⁴

2.2.4 Interpretation of context-dependent expressions

The focus of this dissertation is on the main semantic difference between direct and indirect speech: the interpretation of context-dependent expressions such as the person indexicals *I* and *you*. Before going into detail, imagine the following scenario. Anna tells Peter on Saturday that she will call Ben tomorrow: *Ik ga hem morgen bellen* ‘I will call him tomorrow’. The next day Peter meets Ben and wants to report to him what Anna told him. In Dutch, Peter can choose to do so with a direct speech report like (14a) or an indirect speech report like (14b).

- (14) a. Anna zei: “Ik ga hem morgen bellen”.
 ‘Anna said: “I will call him tomorrow”’
 b. Anna zei dat ze jou vandaag zou gaan bellen.
 ‘Anna said that she would call you today.’

As typical for speech reporting, the scenario involves two speech contexts: the reported context (Anna talking to Peter) and the reporting context (Peter talking to Ben). Both speech contexts are semantically minimally defined by a speaker (and sometimes an addressee), a time and a world (Kaplan, 1989; Schlenker, 2003). Each of these speech contexts can serve as context of evaluation for context-dependent expressions.

In indirect speech, all context-dependent expressions are oriented towards the reporting context with the reporting speaker’s (i.e., Peter’s) time and place of utterance as *deictic center*. This is the reason why in the indirect speech report (14b), the second-person pronoun *jou* ‘you’ refers to Peter’s current addressee Ben and the third-person pronoun *ze* ‘she’ to Anna, who is absent when Peter reports her utterance. Similarly, the temporal adverb *vandaag* ‘today’ has to be evaluated with respect to the actual reporting time, i.e., Sunday.

In direct speech, a context shift from the reporting context to the reported context takes place (Recanati, 2000). Therefore the meaning of the deictic expressions *ik* ‘I’, *hem* ‘him’ and *morgen* ‘tomorrow’ in (14a) has to be determined relative to the reported speech context with Anna’s *I-now-here* as deictic center of the personal and spatio-temporal coordinate system (Bühler, 1934).

⁴ However, note that even though direct speech reports fail the tests of *wh*-extraction and NPI licensing, this does not mean that quotations are semantically and syntactically inert and are completely isolated from the rest of the discourse. Partee (1973) shows that the quoted material can interact in at least three respects with its environment: pronominalization, ellipsis and what she calls ‘semantic anaphora’.

Put differently, we can describe the difference between direct and indirect speech as a difference in perspective: The reporting speaker Peter has the choice to present the target utterance from his own current perspective (indirect speech) or to shift to the perspective of the reported speaker Anna (direct speech) (Coulmas, 1986; De Roeck, 1994).

Now, let's have a closer look at the types of expressions of which the meaning differs when embedded in a direct as compared to an indirect speech report. First of all, this applies to deictic expressions traditionally divided into the semantic categories of person deixis, space deixis and time deixis (Bühler, 1934). This includes deictic pronouns (e.g., *I, you*), spatial indexicals (e.g., *here, there*) and temporal indexicals (e.g., *today, tomorrow*). Second, there are certain types of linguistic expressions that are tied to the speaker's perspective. So-called *expressives* express the speaker's evaluative attitude towards a certain person, object or state of affairs. Examples of words or expressions with expressive content are epithets (e.g., *that bastard*), evaluative adverbs (e.g., *unfortunately, alas!*) and expressive morphemes (*fucking*, German *scheiß-* 'shit-') (Banfield, 1973; Eckardt, 2015; Potts, 2007). Usually, expressives convey the reported speaker's perspective in direct speech and the reporting speaker's perspective in indirect speech, but see section 2.3.4 for exceptions.

2.2.5 Direct speech: A verbatim rendition?

It is an integral part of many theories of reported speech that direct speech – in contrast to indirect speech – commits the speaker to a faithful rendition of the exact words of an original utterance (e.g., Coulmas, 1985; Leech & Short, 1981; Li, 1986). Consider, for instance, Leech and Short's characterization of the direct–indirect distinction:

The essential semantic difference between direct and indirect speech is that when one uses direct speech to report what someone has said one quotes the words used verbatim, whereas in indirect speech one expresses what was said in one's own words. (Leech & Short, 1981, p. 318)

However, this so-called verbatim assumption has also been heavily criticized (e.g., Clark & Gerrig, 1990; Tannen, 2007; Wade & Clark, 1993). Can verbatimness serve as a semantic criterion to distinguish direct and indirect speech?

A major point of criticism against the verbatim assumption is that direct speech reports are usually not verbatim renditions of previous utterances. First of all, many speech reports, such as hypothetical reports (*Daddy would say...*) or fictive reports (*Little Red Riding Hood said...*), are not based on a unique utterance in the past (Tannen, 2007). Second, even if there is an original utterance, direct quotations are often not exact copies of the original, especially in spoken interactions. In an experimental study, Wade and Clark (1993) found that direct speech reports do not contain significantly more exact words than indirect speech reports. They argue that this is due to memory limitations on the part of the speaker.

In addition, the reporting speaker might not always intend to provide a faithful copy of another person's utterance, but may prefer to entertain the listener (Clark & Gerrig, 1990; Wade & Clark, 1993).

However, support for the verbatim assumption is that listeners seem to have the expectation that direct speech reports are verbatim renditions. Psycholinguistic studies indicate that people pay more attention to the exact words of a direct than an indirect speech report (Bohan, Sanford, Cochrane, & Sanford, 2008; Eerland et al., 2013). Moreover, speakers tend to be willing to admit error when confronted with written or recorded evidence that their direct quote of someone is inaccurate (Johnson & Lepore, 2011).

In sum, verbatim reports seem to be more a theoretical ideal than actual reality in spoken conversations. Verbatimness therefore falls short of being an empirically testable criterion to distinguish direct from indirect speech. But even though speakers might *de facto* not quote the exact words of another person, they purport to do so by presenting the reported utterance not from their own perspective but from the perspective of the reported speaker.

2.2.6 Prosodic differences

In languages like English, syntactic features are not always sufficient to determine whether a report is direct or indirect speech. Consider, for instance, report (15), which is ambiguous between a direct and an indirect speech interpretation: The first-person pronoun *I* could either refer to the reported speaker Anna (direct speech) or to the person producing the speech report (indirect speech).

(15) Anna said I am lucky to be alive.

In spoken language, prosody is an important means to disambiguate reports like (15). Studies indicate that direct speech reports tend to exhibit a break between reporting clause and report and are characterized by a greater overall pitch range than indirect speech reports. Speakers also sometimes change their volume, rhythm, speech rate and voice quality in a direct quotation (Jansen, Gregory, & Brenier, 2001; Kalmanovitch, 2015; Kasimir, 2008; Klewitz & Couper-Kuhlen, 1999; Oliveira & Cunha, 2004; Wade & Clark, 1993).

Taken together, indirect speech reports tend to have an integrated intonational contour, whereas direct quotations are prosodically separated from the surrounding discourse. The prosodic discontinuity in direct speech is a cue for the listener that a shift from the reporting to the reported speech context takes place. However, note that there is a lot of variation in the prosodic marking of speech reports. Therefore, prosody is not always a clear indicator of direct or indirect speech.

In face-to-face interactions, the context shift in direct speech can also be signaled visually, e.g., by air quotes or by mimicking the reported speaker's facial expression, posture or physique (Clark & Gerrig, 1990). In written language, quoted speech is usually distinguished from the surrounding text by means of quotation marks, italics, indentation or different font size or type.

In the following sections, I discuss two pragmatic differences between direct and indirect speech. First, direct speech is usually perceived as more vivid than indirect speech (2.2.7). Second, there is evidence that direct and indirect speech are used for different communicative purposes (2.2.8).

2.2.7 The vividness of direct speech

Pragmatically, direct and indirect speech can be distinguished with respect to the vividness of the report. Direct speech is typically described as “vivid”, “dramatic” and “theatrical” and is considered to be more engaging and involving for an audience than indirect speech reports (Chafe, 1982; Tannen, 2007; Wierzbicka, 1974). It is empirically confirmed that listeners perceive direct speech as more lively than other stretches of talk (Groenewold, Bastiaanse, Nickels, & Huiskes, 2014). This is related to certain features of direct speech mentioned earlier: the frequent presence of expressive elements, the prosodic changes in pitch, volume, tempo and voice quality, and the shift to the reported speaker's perspective. The vividness distinction can be explained by the difference between demonstration and description, proposed by Clark and Gerrig (1990).

Demonstrations are inherently more vivid than descriptions [...]. As demonstrations, direct quotations are intended to enable listeners to experience what it would be like to hear, see, or feel what the original speaker did. Indirect quotations, which are descriptions, are not. (Wade & Clark, 1993, p. 818)

Compared to other differences between direct and indirect speech, the vividness criterion has received a lot of attention in psycholinguistics and neurolinguistics. Several studies indicate that the alleged vividness of direct speech affects listeners' and readers' mental representations. There is experimental evidence that readers adapt their reading rate in direct speech but not in indirect speech when the situation implies that the reported speaker is speaking quickly or slowly (Stites, Luke, & Christianson, 2013; Yao & Scheepers, 2011). Moreover, reading direct as opposed to indirect speech reports activates specific voice-selective areas in the brain, which suggests that the reported speaker's voice is simulated (Yao, Belin, & Scheepers, 2011). A similar effect was found when participants listened to monotonously spoken direct and indirect speech:

The current study shows that listeners routinely expect vivid depictions for direct speech but rarely for indirect speech; they spontaneously engage in men-

tal simulations of vivid vocal depictions while listening to monotonously spoken direct speech rather than to monotonously spoken indirect speech. (Yao, Belin, & Scheepers, 2012, p. 1841)

The theoretical and empirical evidence suggests that vividness is a feature typically associated with direct speech. But as a criterion to distinguish direct and indirect speech, vividness has some serious disadvantages. First, vividness is a subjective feature related to how people perceive a speech report. Second, vividness is a gradual feature, which means that certain direct speech reports are very vivid while others might exhibit no expressive elements or exaggerated prosody at all, similar to indirect speech (Redeker, 1991).

2.2.8 Different discourse functions

Even though direct and indirect speech differ in many syntactic, semantic and prosodic respects, there is no “one-to-one relationship between direct (or indirect) speech and a particular context-free function” (Günthner, 1997, p. 268). Instead, conversation analysts argue that the function of a particular direct or indirect speech report needs to be analyzed with respect to the concrete situation of its use (Baynham, 1996; Günthner, 1997).

Here are a few examples of what people do with direct and indirect speech reports. Speakers tend to use direct speech to create a dramatic effect, which makes the speech report more involving for the listener (Baynham, 1996; Redeker, 1991; Sakita, 2002; Tannen, 2007; Wierzbicka, 1974). This is why direct speech is not only frequently used in narratives, but for instance also by teachers during math class (Baynham, 1996). Within a story or joke, direct speech often appears at the climax, whereas indirect speech reports have the function to contribute background information (Bauman, 1986; Günthner, 1997; Sakita, 2002; Yule, 1993). In court, witnesses can take advantage of the fact that direct speech is associated with verbatimness. Using a direct instead of an indirect speech report gives the impression that a person has directly witnessed the speech act in question and has a detailed memory of it. This can strengthen the credibility of the testimony (Galatolo, 2007; Philips, 1986).

While there is a multitude of qualitative studies, quantitative studies about the discourse functions of direct and indirect speech are rare. Vincent and Perrin’s (1999) analysis is a notable exception. Their study of speech reports produced during sociolinguistic interviews reveals interesting functional differences between direct and indirect speech. While the majority of direct speech reports (54%) have a narrative function, only 34% of indirect speech reports are used to move along a story chronologically. The most frequent function among indirect speech reports is the authority function. Lucy’s report at the outset is an example of a report with an authority function – here repeated for convenience:

- (1) My mom said that I am not allowed to swim in the river.

Lucy uses report (1) to communicate the proposition that she is not allowed to swim in the river, which she supports by appealing to the authority of her mother. That Lucy in our fictive example opted for an indirect speech report is plausible considering that the authority function is associated with the use of indirect speech. In Vincent and Perrin's corpus, 36% of the indirect speech reports have an authority function as opposed to 16% of the direct speech reports.⁵

In sum, direct and indirect speech cannot be clearly distinguished with respect to their function in discourse. Both report types can be used for multiple communicative purposes. However, in certain situations, speakers seem to prefer one type of report over the other. To gain a better understanding of the factors that influence speakers' choice of direct or indirect speech, experimental studies are needed in which the situational parameters are systematically controlled and varied. I provide such experimental evidence in the elicitation study in chapter 5.

2.3 BETWEEN DIRECT AND INDIRECT SPEECH

We have seen how to distinguish direct and indirect speech by various syntactic, semantic and prosodic features. In Dutch, a language with an especially clear-cut direct–indirect distinction, direct and indirect speech can be distinguished by word order (V2 vs. V-final), the presence/absence of a complementizer, the occurrence of certain words or constructions in the report (e.g., vocatives, imperatives), the performance on syntactic and semantic tests (*wh*-extraction, NPI-licensing), the deictic orientation of context-dependent expressions (reported vs. reporting context), and particular prosodic features. Furthermore, direct and indirect speech differ also in pragmatic respects: Direct speech reports are usually perceived as more vivid than indirect speech reports and tend to be used for different communicative functions. These differences between direct and indirect speech support the traditional dichotomy in which languages provide two entirely distinct linguistic structures to report what someone said: one a form of quotation and the other an intensional embedding. However, on closer examination, it turns out that even in Dutch some reports do not neatly fit into the dichotomy.

⁵ Note that in Vincent and Perrin's (1999) corpus, 87% of the reports were direct speech reports and 13% indirect speech reports. This is in line with previous studies that found that direct speech is the most frequent report type in spoken discourse, whereas indirect speech dominates in written newspaper texts (Barbieri & Eckhardt, 2007; Van der Houwen, 2012). Beside the narrative function and the authority function, Vincent and Perrin distinguish two more functions of speech reports: the support function and the appreciative function. A report has a support function when it illustrates a meta-discursive comment by the speaker. A report has an appreciative function when it expresses an opinion or judgement about something. Vincent and Perrin found that the support function is associated with direct speech, while the appreciative function is associated with indirect speech.

In the next sections, I will go through apparent counterexamples to a rigid direct–indirect distinction in Dutch: V2 constructions, mixed quotation, free indirect discourse, and perspective-dependent expressions. Looking beyond Dutch, many more direct–indirect mixes turn up, which will lead us to endorse a more continuous picture of reported speech.

2.3.1 V2 constructions

A first piece of evidence that the direct–indirect distinction in Dutch is not as strict as it appeared until now is that in colloquial Dutch, we find an embedded verb-second construction that is similar to the German indirect V2 construction:

- (16) Jan zei (dat) hij kon niet komen.
 ‘John said (that) he couldn’t come.’ (Zwart, 1997, p. 24)

This construction is situated in between direct and indirect speech because it exhibits the pronoun interpretation and sequence-of-tense characteristics of indirect speech, but the word-order (and *wh*-extraction blocking) of direct speech. Although not acceptable in standard Dutch, verb-second mixed reporting is considered grammatical in many Germanic languages, including German, Danish, and Frisian (cf. Zwart, 1997 and references cited therein).⁶

2.3.2 Mixed quotation

Mixed quotations such as (17) also contradict the idea of a clear-cut dichotomy because they combine features of direct and indirect speech (Cappelen & Lepore, 1997).

- (17) Peter zei dat de “flamingo” een mooie dans is.
 ‘Peter said that the “flamingo” is a beautiful dance.’

In report (17), the reporting speaker uses the syntax of indirect speech, but directly quotes Peter’s malapropism “flamingo”. According to Davidson (1979, p. 39), the quoted words in mixed quotations “do double duty” because they are simultaneously mentioned (i.e., refer to a linguistic object) and used (i.e., fulfill a syntactic function and refer to entities in the world).⁷ Since mixed quotations are a frequent phenomenon in spoken and written language, it is hard to ignore this evidence against a rigid direct–indirect distinction.

⁶ There is a similar phenomenon in some English dialects, where the main clause word order can be seen in reported questions, as in: “The baritone was asked what did he think of Mrs Kearney’s conduct.” (from James Joyce, *Dubliners*, cited by McCloskey (2006)).

⁷ In the formal semantic analysis of Maier (2014), example (17) would roughly be analyzed as: Peter said that the entity he refers to with the word *flamingo* has the property of being a beautiful dance.

2.3.3 Free indirect discourse

Another well-known counterexample to the direct–indirect dichotomy is free indirect discourse, a narrative technique typically used to report the thoughts of a character (Banfield, 1982; Fludernik, 1993; Leech & Short, 1981). But free indirect discourse can also be used to report speech, as in (18).

- (18) Ze keek hem woest aan. Wie dacht hij wel dat hij was?!, snauwde ze.
‘She looked at him furiously. Who did he think he was?!, she snarled.’

In free indirect discourse, the perspective of the narrator (i.e., the reporting speaker) and the perspective of the character (i.e., the reported speaker) are intermingled. On the one hand, (18) looks like a direct report of the character’s speech because of the use of an exclamative construction with main clause word order. On the other hand, the tense and the use of a third-person pronoun *hij* ‘he’ point in the direction of indirect speech, as the character’s original utterance must have been something like *Who do you think you are?!*. Because free indirect discourse mixes features of direct and indirect speech, it has not only attracted the attention of literary scholars, but also that of formal semanticists (e.g., Eckardt, 2015; Maier, 2015a; Schlenker, 2004; Sharvit, 2008).

2.3.4 Perspective-dependent expressions

As noted in section 2.2.4, the main semantic criterion to distinguish direct and indirect speech is the difference in perspective from which the utterance is reported: In indirect speech, the utterance is presented from the reporting speaker’s perspective, and in direct speech from the reported speaker’s perspective. However, not all linguistic expressions behave as predicted by this distinction. Consider, for instance, the meaning of the speaker adverb *helaas* ‘unfortunately’ in (19).

- (19) Anna zei dat ze helaas niet kon komen, maar ik vind het niet zo erg.
‘Anna said that she unfortunately could not come, but I think that is not so bad.’

Since (19) is an indirect speech report, the speaker-oriented adverb *helaas* should express the actual reporting speaker’s attitude of regret. However, the linguistic context rules out such an interpretation because the actual speaker is clearly not unhappy that Anna cannot come (*maar ik vind het niet zo erg* ‘but I think that is not so bad’). Contrary to our initial characterization, *helaas* expresses the perspective of the reported speaker Anna, similar to when it occurs in a direct speech report (cf. Eckardt, 2015; Fabricius-Hansen, 2002 for *leider* ‘unfortunately’ in German). Exceptions like this have also been discovered for epithets (e.g., *that bastard*) and appositives (Harris & Potts, 2009; Koev, 2013; Kratzer, 1999).

Even temporal and spatial indexicals are not always firmly anchored in either the reported context (direct speech) or the reporting context (indirect speech). Plank's (1986) study indicates that native speakers of German allow deviant interpretations of the spatial adverbs *hier* 'here' and *dort* 'there' and the temporal adverbs *heute* 'today' and *morgen* 'tomorrow' in both direct and indirect speech. A similar study for Dutch is still missing, but the data from German suggests that in Dutch the direct–indirect distinction might also be blurred in the domain of space and time deixis. Crucially, however, in the domain of person deixis, Dutch and German do not allow pronouns in direct or indirect speech to be evaluated with respect to the opposite speech context. This makes pronouns an ideal test case for finding out when children can reliably tell apart direct and indirect speech.

2.3.5 Role shift in sign languages

If we look beyond Dutch, we find many more mixes of direct and indirect speech. Consider first sign languages, where utterances are reported with a construction called *role shift*. In role shift, the shift to another person's perspective is marked in the visual-gestural modality and is characterized by a body shift toward the locus of the quoted speaker, break in eye gaze with the actual addressee, change in head position and the use of facial expressions associated with the quoted speaker (Engberg-Pedersen, 1995; Herrmann & Steinbach, 2012; Quer, 2005, 2011). Role shift is often considered a form of direct speech because of its demonstrative qualities and the fact that deictic expressions are typically evaluated with respect to the reported signer's perspective (Davidson, to appear).

However, on closer inspection, a simple identification of role shift with direct speech seems premature. In several sign languages such as Catalan Sign Language (Quer, 2005) and German Sign Language (Herrmann & Steinbach, 2007; Hübl, 2013), particular spatial indexicals (e.g., *here*) and temporal indexicals (e.g., *now*) have a preference not to shift under role shift. This means that they can be anchored in the actual reporting context, like in indirect speech. What is more, in Danish Sign Language even person indexicals can have an unshifted interpretation in role shift, as example (20) indicates.

- (20) My mother_i told me_j; “[...] I_j shall stay in Nyborg”. (Engberg-Pedersen, 1995, p. 139, translation from Danish Sign Language)

In example (20), the reporting signer refers to herself with a first-person pronoun inside the role shift quotation of her mother's speech. According to Engberg-Pedersen (1993), it is not even possible for the reporting signer to use a second- or third-person pronoun for self-reference in the context of role shift. This is the reason why in Danish Sign Language, utterances are usually reported with indirect speech reports when they include reference to people present in the actual speech context.

2.3.6 Person magnetism

Interestingly, cases of unshifting in a supposedly direct speech report have also been observed in the spoken languages Kwaza (Brazil), Slave (Canada), and Nez Perce (United States) (Evans, 2012). Kwaza, for instance, has a report type that displays all features of canonical direct speech, but with one notable exception concerning the quoted second person singular. In the Kwaza equivalent of (21), the interpretation of *you* is ambiguous.

(21) Margarida says, “You are ill”.

You can refer to the addressee of the reported context, who is usually identical with the reporting speaker. This is the expected interpretation of second-person pronouns in direct speech and means that Margarida’s original utterance must have been something like “You are ill”. An alternative interpretation of *you* is available when the referential target person is the addressee in the reporting context (Original \approx “He/she is ill”). Under this condition, second-person *you* can have an unshifted meaning, like in indirect speech (Van der Voort, 2004). While this kind of unquotation occurs only with second-person pronouns in Kwaza and Slave, first-person pronouns are affected as well in Nez Perce (Evans, 2012).

Evans (2012) calls this phenomenon “speech act participant attraction” or “person magnetism”, suggesting that the addressee and the speaker of the actual reporting speech context are so salient that they ‘attract’ pronouns. As we have seen above, in sign languages spatial indexicals such as *here* can also be ‘attracted’ by the actual speech context. We can therefore more generally speak of an *actual speech context attraction*, which can affect deictic expressions in a (seemingly) direct speech report.

2.3.7 Towards a more fluid picture

Taken together, the examples from Dutch and other languages have illustrated that reports can simultaneously exhibit features of direct and indirect speech. This makes a clear-cut split into direct and indirect speech reports impossible. While I could only mention a few examples of direct–indirect mixes above, there are many more ways in which languages can and do combine features of direct and indirect speech:

Once things get more mixed the number of possibilities is astronomical: the full possibility space (impossible to show here) would be the product of all dimensions (person, tense, mood, honorificity, space, evaluation, etc.), times all coding sites (e.g. subject, object, possessor, etc.), times all values (e.g. first vs. second vs. third person). (Evans, 2012, p. 98)

De Roeck, who analyzed speech reports in a sample of 40 languages, comes to the conclusion that the question whether direct and indirect speech “form a dichotomy definitely calls for a negative answer” (De Roeck, 1994, p. 346). Haberland, who provides a thorough study of reported speech in Danish, suggests to view direct and indirect speech as “two tendencies” rather than “two mutually exclusive and clearly distinguishable sets” (Haberland, 1986, p. 248). Ebert sums up her study of three Nepalese languages by stating that the “idea that languages make a clear distinction between direct and indirect speech is for the most part a grammatical fiction” (Ebert, 1986, p. 156).

The abundance of cross-linguistic direct–indirect mixes leads some theorists to abandon the strict direct–indirect dichotomy in favor of a more fluid picture. Evans (2012) proposes a canonical approach with three types of speech reports: canonical direct speech, canonical indirect speech, and canonical biperspectival speech. In these idealized types, all semantic values in all domains express the same perspective: either the reported speaker’s perspective (canonical direct speech), the reporting speaker’s perspective (canonical indirect speech), or simultaneously both the reporting and the reported speaker’s perspective (canonical biperspectival speech).⁸ Crucially, these three canonical types are merely ideals spanning a continuum of possible mixed forms. While canonical direct speech is attested in many languages, canonical indirect speech is relatively rare, and canonical biperspectival speech even unheard of in any natural language – probably due to the cognitive complexity of simultaneously taking into account two perspectives (Evans, 2012). The benefit of stipulating ideals is to sketch the complete space of possible reports. While Evans focuses on features situated in the area of semantics, his framework could be extended to include also features related to the degree of syntactic integration (e.g., word order, complementizer), syntactic/semantic transparency (*wh*-extraction, licensing of negative polarity items), and possibly also prosodic and pragmatic aspects.

Maier (2009) also argues against a rigid direct–indirect distinction. Within the framework of formal semantics, he proposes a unified treatment of speech reports, using the semantic mechanism of mixed quotation (Geurts & Maier, 2005). In Maier’s account, speech reports are essentially treated as indirect reports, in which parts of the reported utterance can be quoted, i.e., presented from the reported speaker’s perspective. Canonical direct and canonical indirect speech then turn out to be the limiting cases on a continuum of mixed reports. Direct speech means that the complete reported utterance is quoted and indirect speech that no quotation is involved.

⁸ According to Evans (2012), logophoric pronouns are examples of biperspectival expressions. The logophoric pronoun *yè*- in the West African language Ewe, for instance, indicates that its referent is a third person in the reporting context and the speaker in the reported context. Furthermore, complex tenses in indirect speech (e.g., *Peter said that he would go home*) can be analyzed as biperspectival since they depend on two temporal reference points: one derived from the reporting context, the other from the reported context.

Section 2.3 has revealed that a simple direct–indirect dichotomy cannot capture the vast variety of speech reporting constructions that are available in languages all over the world. However, direct and indirect speech (and the features that characterize them) can serve as idealized points to classify and locate speech reports on a continuous scale. Dutch is a language that has grammaticalized two report types that are relatively close to the canonical ideals of direct and indirect speech. For my three experimental studies (reported in chapters 4, 5 and 6) it is crucial that Dutch does not allow deviant interpretations of pronouns in direct and indirect speech. This makes it possible to find out at what age children are able to correctly identify the referents of pronouns embedded in direct and indirect speech. As we will see, even eleven-year-old Dutch children sometimes interpret pronouns in direct speech like in indirect speech when the actual speech context is highly salient. This bears a striking resemblance to the findings from Kwaza and Danish Sign Language discussed above.

2.4 CHILDREN'S ACQUISITION OF DIRECT AND INDIRECT SPEECH

The search for criteria to distinguish direct and indirect speech in section 2.2 has shown that speech reports are a complex linguistic phenomenon related to all levels of linguistic analysis. This implies that children need to combine syntactic, semantic, prosodic and pragmatic knowledge in order to produce and comprehend speech reports in an adult-like fashion. In languages with two types of speech reports, such as Dutch and German, children have to learn which clusters of features are associated with direct and which with indirect speech.

In the following, I will present previous research on children's acquisition of direct and indirect speech. I start with a general overview of children's development of speech reports, before I highlight more specific aspects related to children's production and comprehension of reported speech. The picture that will emerge is that children have a less rigid direct–indirect distinction than adults. This is evident from the fact that children tend to mix direct and indirect speech.

2.4.1 General overview of children's acquisition of speech reports

Research on children's production of speech reports tends to be based on either elicited narratives or spontaneous speech. The majority of studies has focused on children's production of speech reports in elicited narratives (Ely & McCabe, 1993; Emmorey & Reilly, 1998; Goodell & Sachs, 1992; Hickmann, 1993; Nordqvist, 1998; Özyürek, 1996). However, there are also several studies on children's spontaneous production of direct and indirect speech in English (Ely & McCabe, 1993), Swedish (Nordqvist, 2001a, 2001b) and the sign languages ASL and Libras (Lillo-Martin & de Quadros, 2011).

These previous studies reveal that children acquiring English and Swedish start to report speech at around two years of age (Ely & McCabe, 1993; Nordqvist, 2001a). Children acquiring a sign language produce their first speech reports probably at around a similar age

or slightly later (Emmorey & Reilly, 1998; Lillo-Martin & de Quadros, 2011). In the following, I focus mainly on the development of children who acquire a spoken language because my empirical studies will also deal with the acquisition of direct and indirect speech in the spoken languages Dutch and German.

Children's early speech reports are syntactically reduced and consist only of a direct quotation without a metalinguistic frame such as *Anna said* (Nordqvist, 2001a). In this early phase, the interlocutors of the child play an important role in establishing whose speech is being reported. By prompting speech reports and reinterpreting children's utterances, they compensate for the lack of a reporting clause. Nordqvist gives an example of such a "scaffolded" report from an interaction between the Swedish girl Tea (age 2;2) and her mother:

- (22) MOT: what do you say then?
 TEA: hurts
 MOT: do you say that it hurts?
 (Nordqvist 2001a, p. 155, translated from Swedish)

Children's first speech reports with a reporting clause are direct quotations. It is a potentially universal trend that direct speech is acquired before indirect speech, as documented for English (Ely & McCabe, 1993), Swedish (Nordqvist, 2001a) and Turkish (Özyürek, 1996). This finding is not surprising if we consider that the form of the quoted utterance is syntactically less constrained in direct than in indirect speech. Since even single words (e.g., *says "hello"*) or sounds (e.g., *says "moo"*) can be quoted, direct speech reports can be used by young children with only limited syntactic abilities. By contrast, indirect speech constructions tend to be syntactically more complex. In Dutch, for instance, the reported utterance needs to be syntactically integrated as subordinate clause with verb-final word order and complementizer. If we want to guarantee a fair comparison between direct and indirect speech, we should compare direct quotations of complete sentences with indirect speech reports. In my corpus study in chapter 3, I will therefore distinguish between clausal and non-clausal direct speech.

2.4.2 Acquisition of speech reports in Dutch and German

So far, children's development of direct and indirect speech in Dutch and German has not been studied. However, previous studies on the acquisition of complementation in these languages can provide valuable insights because in both Dutch and German, indirect speech is realized in the form of a complement clause dependent on a matrix clause with a verb of saying. As pointed out in section 2.2.1, both Dutch and German have an indirect speech construction with verb-final word order and complementizer. Dutch children begin to use embedded clauses with verb-final word order between 3 and 3;6 years (Bol & Kuiken, 1988).

German children can make use of an additional indirect speech construction, which has verb-second word order and lacks a complementizer. This raises the question whether they acquire the indirect verb-final or the indirect verb-second construction first. According to Rothweiler (1993), German children's first embedded clauses contain verbs in clause-final position. If this general statement is correct, it would imply that German children acquire indirect speech with verb-final word order before indirect speech with verb-second word order. However, based on a large corpus of spontaneous speech, Brandt, Lieven and Tomasello (2010) show that neither the verb-final construction nor the verb-second construction is acquired first. German children start using complements with both word order patterns from around two years of age, but with different sets of complement-taking verbs.

The German boy Leo, whose language development between 2;0 and 5;0 has been meticulously documented, uses, for instance, the complement-taking verb *wissen* 'know' primarily with verb-final complements, but *glauben* 'believe' with verb-second complements. Most relevant for the acquisition of speech reports are complement-taking communication verbs. Both Leo and his mother have a clear preference for using *sagen* 'say' with verb-second complements (Leo: 70%, Mother: 82%) and *fragen* 'ask' with verb-final complements (Leo: 76%, Mother: 100%) (Brandt et al., 2010).

However, caution is required when interpreting these findings. Brandt et al. also counted direct quotations of assertions as verb-second complements. This comes to light when we have a look at their example of a verb-second complement, reproduced in (23).

- (23) CHI: Mama soll sagen, du musst nicht husten.
 'Mum should say, you don't have to cough.'
 MOT: Du musst nicht husten, Leo.
 'You don't have to cough, Leo.' (Brandt et al., 2010, p. 601)

In the conversation with his mother, Leo (age: 3;0) uses the report *Mama soll sagen, du musst nicht husten* 'Mum should say, you don't have to cough'. That this is a direct speech report is evident from the shifted use of *du* 'you' which refers to Leo himself and not to Leo's current addressee, his mother. The mother's subsequent response confirms this interpretation. This indicates that Brandt et al. collapsed direct speech and verb-second indirect speech – two syntactically and semantically very different types of speech reports – in their analysis. Therefore the insights that we can gain from their study about the frequency of different report constructions with the *verbum dicendi sagen* are quite limited. In chapter 3, I present my own corpus analysis of German children's spontaneous speech reports, in which I draw the necessary distinctions between (clausal and non-clausal) direct speech, verb-second indirect speech and verb-final indirect speech.

After this general overview of children's syntactic development, I will now highlight two aspects in which children's speech reports differ from those of adults: the lack of clearly marked perspective shifts (2.4.3) and mixes of direct and indirect speech (2.4.4).

2.4.3 Marking of perspective shifts

A stable cross-linguistic finding is that the direct speech reports of younger children often lack a reporting frame (Goodell & Sachs, 1992; Hickmann, 1993; Nordqvist, 2001a; Özyürek, 1996). This makes it difficult or impossible for a listener to decide whose perspective a certain utterance expresses: that of the actual speaker or that of someone else. This effect is aggravated by the fact that younger children rarely change their pitch or voice quality in unframed quotations (Goodell & Sachs, 1992). Similarly, children acquiring a sign language often do not explicitly convey whose (speech) act they present in role shift and do not consistently use shifted facial expressions as a role shift marker (Emmorey & Reilly, 1998; Lillo-Martin & de Quadros, 2011).

Nordqvist's (2001a) analysis of Swedish children's frog story narratives⁹ can give an impression how frequent reports with poorly signaled perspectives are. In the oral narratives of three-year-old children, 49% of the speech reports did not indicate who the reported speaker was. This high number dropped to 34% in four-year-olds and further decreased until the age of twelve, when all kids marked perspective shifts successfully (Nordqvist, 2001a). In written texts, the problem continues even longer. Around 20% of nine- and twelve-year-old's written speech reports left the reader clueless about the identity of the reported speaker (Nordqvist, 2001a). This could be due to the fact that in writing prosodic modifications such as changes of pitch or voice quality are not available to mark perspective shifts, while at the same time quotation marks are acquired relatively late (see section 2.4.8).

The finding that children often do not explicitly mention whose speech they report suggests that they fail to take the addressee's perspective into account. This failure is not specific to speech reporting, but has also been observed in other linguistic areas, for instance in the production of referring expressions (e.g., Deutsch & Pechmann, 1982; Hendriks, Koster, & Hoeks, 2014).

2.4.4 Direct–indirect mixes

A second characteristic of children's speech reports is that they sometimes include features of both direct and indirect speech (Goodell & Sachs, 1992; Hickmann, 1993; Nordqvist, 2001a).

In Goodell and Sachs's study (1992), children were asked to retell conversations between Tommy Inchworm and several birds. In the original text, the bird says to Tommy Inchworm: "Thanks for measuring my beak. I feel more beautiful already". An eight-year-old child reports this utterance as follows:

⁹ Nordqvist uses the wordless picture book "Frog, where are you?" by Mercer Mayer to elicit narratives from children. This book has been used to examine children's production of narratives in many languages (e.g., Berman & Slobin, 1994).

- (24) And the birdie_i said that I_i feel much prettier knowing how long my_i beak is.
(age: 8;6) (Goodell & Sachs, 1992, p. 407)

This report contains on the one hand signs of direct speech, such as the shifted pronouns *I* and *my*, presented from the reported speaker's perspective. On the other hand, (24) contains the complementizer *that*, which suggests that it is an indirect speech report.

Direct–indirect mixes like this are frequent in children's elicited retellings of stories. Goodell and Sachs (1992) found that 22% of four-year-old children's 'indirect' speech reports were direct–indirect mixes. The frequency of these mixed speech reports decreased with age to 16% in six-year-olds, 4% in eight-year-olds and 0.5% in adults. The largest number of direct–indirect mixes is related to inconsistencies in pronoun use, like in example (24). Another frequent type of mix concerns sequence of tense. Children either fail to apply sequence of tense rules in indirect speech or incorrectly backshift the verb in direct speech, as in (25).

- (25) Then he said, "How dare you say I was ugly?" (age: 4;3)
(Goodell & Sachs, 1992, p. 407)

These examples indicate that children have difficulties to clearly distinguish between direct and indirect speech. Children's lack of a rigid direct–indirect distinction results in the production of reports that contain features of both report types. The question arises whether children have similar problems in the comprehension of speech reports. In the following sections, I review previous studies on children's comprehension of direct and indirect speech. In particular, I look at children's comprehension of complement clauses (2.4.5) and their interpretation of *wh*-sentences (2.4.6).

2.4.5 Children's comprehension of complement clauses

Children's understanding of direct and indirect speech reports has not been directly compared yet. However, courtroom studies suggest that children do not always process embedded sentences in an adult-like way (Carter, Bottoms, & Levine, 1996; Walker, 1993). Consider the following dialogue taken from a psychological interview with a five-year-old child that took place before trial:

- (26) Q: Do you remember when Don asked you, "What color was their skin, like mine or like Martha's?"
CHI: Like yours (Walker, 1993, pp. 70–71)

The question asked by the interrogator contains two levels of embedding: [Do you remember when [₁ Don asked you, [₂ "Y"]]]. The child effectively ignored the outer layers and instead answered the quoted question ascribed to Don about the skin color of certain people.

Experimental studies confirm the observation that children's comprehension of complement clauses differs from that of adults. Jill de Villiers and collaborators (De Villiers & De Villiers, 2000; De Villiers & Pyers, 2002) found that children below the age of four have difficulties understanding sentences in which the embedded clause expresses a false proposition. They tested three- and four-year-old children with stories like (27), that were presented together with pictures.

- (27) The Mom said she bought apples, but look, she really bought oranges. What did the Mom say she bought? (De Villiers, 2005, p. 186)

In the situation described in (27), what the mother said she bought (apples) and what she really bought (oranges) differs. This means that the mother intentionally or unintentionally misrepresents her purchases. When children had to answer the embedded question *What did the Mom say she bought?*, four-year-olds gave the correct answer *apples*, whereas three-year-olds incorrectly answered *oranges*. This suggests that young children fail to integrate the complement under the scope of the verb *say* and give a reality-oriented answer instead.

There is evidence from various correlational and training studies (De Villiers & De Villiers, 2000; De Villiers & Pyers, 2002; De Villiers & De Villiers, 2012; Hale & Tager-Flusberg, 2003; Lohmann & Tomasello, 2003; Tager-Flusberg & Joseph, 2005; De Villiers, Hobbs, & Hollebrandse, 2014) that children's acquisition of complement structures is related to their theory of mind development. De Villiers and colleagues even claim that the syntax of complementation (i.e., the syntax of indirect speech or thought reports) is a necessary prerequisite for theory of mind development because it provides the mental format for representing false beliefs (De Villiers & De Villiers, 2000).¹⁰

¹⁰ Ruffman, Slade, Rowlandson, Rumsey, and Garnham (2003) question this interpretation. They argue that De Villiers's methodology confounds children's comprehension of sentential complementation with false belief understanding: Only children who already have a basic understanding of false belief are able to memorize mistaken propositions in scenarios like (27). De Mulder (2011) who took Ruffman et al.'s (2003) methodological concerns into account and used a revised version of the memory for complements test, subsequently found no correlation between children's understanding of complement clauses and their theory of mind performance. An alternative to De Villiers's linguistic determinism has been proposed by Perner et al. (2003). They argue that the determining factor for children's theory of mind development is conceptual progress rather than the acquisition of a certain linguistic structure. They provide evidence from German, where the complement-taking verbs *wollen* 'want', *sagen* 'say' and *denken* 'think' are all realized with the same syntactic construction. Nevertheless, German children understood *want that* before *say that* and *think that*, arguably because of a conceptual desire-belief gap (Perner, Sprung, et al., 2003).

Surprisingly, what has been neglected in this debate is that direct quotations exhibit similar features as the indirect speech or thought reports that are used in the complement tests. In direct speech reports like (28), the truth value of the complete sentence is also independent of the truth value of the quoted sentence.

(28) The Mom said, “I bought apples”.

Interestingly, Tager-Flusberg and Joseph (2005) actually used direct speech and thought quotations when testing autistic children on a ‘complement’ test, arguing that these direct reports are linguistically less complex and therefore easier to understand for their participants. They found that the ability to understand speech quotations that express false propositions was the crucial predictor for theory of mind in children with autism.

This indicates that not the syntax of complementation is the crucial factor, but the fact that speech or thought reports are metalinguistic utterances with two levels of content: an object-representation and a “‘meta’ part, whereby the object-representation is referred to as entity on its own right and situated in the order of things” (Recanati, 2000, p. xii). If these considerations are correct, children should be able to understand direct and indirect speech reports that contain a false embedded statement at the same age. Additional correlational and training studies could clarify whether the comprehension of direct speech reports is also related to children’s false belief understanding.

2.4.6 *Wh*-movement

As pointed out in section 2.2.3, *wh*-movement has been proposed as a test to distinguish direct from indirect speech reports. Languages such as English, German or Dutch only allow *wh*-extraction out of indirect speech reports, but not out of quotations. However, in contrast to adults, children do not always obey the *wh*-movement block in quotations (Hollebrandse, 2007; Weissenborn, Roeper, & De Villiers, 1991).

In Hollebrandse’s (2007) experiment, three- to seven-year-old English-speaking children listened to stories like the following: Bart, the experimenter, and Deanne are going for a bike ride. Bart wears a helmet and Deanne wears blue gloves. But because they are not sure whether this is actually safe enough, Deanne calls Daddy at work and asks him whether she can ride a bike with her blue gloves and Bart with a helmet. Experimenter Bart then asks the participant the following test question, which contains a short pause before the quote:

(29) How did Deanne ask, “Can I ride a bike?”

The preceding story context provides three possible answers: (a) “by phone” (direct: *I* referring to reported speaker Deanne, no *wh*-extraction), (b) “with a helmet” (indirect: *I* referring to reporting speaker Bart, *wh*-extraction), (c) “with blue gloves” (mix: *I* referring to reported

speaker Deanne, *wh*-extraction). The majority of participants gave the correct short-distance answer (a), but quite a few participants also chose the long-distance answers (b) or (c). While three- and four-year-olds showed no preference for option (b) or (c), five- and six-year-olds clearly preferred (c). This means that they interpreted *I* as referring to Deanne, like in direct speech, but still allowed *wh*-movement out of the complement, which indicates indirect speech. This suggests that these children mix features of direct and indirect speech in their interpretation of speech reports.

Note that Hollebrandse (2007) himself draws a different, much stronger conclusion: Children can extract out of quotations. This presupposes that children recognized that they were dealing with direct rather than indirect speech. According to Hollebrandse, the shifted interpretation of the pronoun establishes that children interpreted the sentence as direct speech. But one could just as easily argue that the *wh*-movement establishes an indirect speech interpretation, so that the data show that children allow context shifting in indirect speech. In fact, this is precisely the type of argument that Schlenker (2003) and Anand (2006) use for introducing context shifters in indirect speech in languages like Amharic. Maier's (2009) and Evans' (2012) continuous account of speech reports sidesteps these issues. Instead of clinging to any one characteristic as showing that a given report is either direct or indirect, they allow the possibility of mixed reports that exhibit characteristics of both direct and indirect speech.

Spoken and written language differ in the way quotations are marked. In spoken language, quotations can be signaled prosodically with a break between reporting clause and quote, and a change of pitch or voice quality in the quotation. In written language, punctuation marks indicate the deictic shift to another person's perspective. In the following sections, I have a look at children's acquisition of quotation marking in the spoken (2.4.7) and written modality (2.4.8).

2.4.7 Prosodic quotation marking

Children start early with marking direct quotations prosodically. In her analysis of frog story narratives, Nordqvist (2001a) found that most three-year-old and all four-year-old Swedish children separated reporting clause and quotation by a short pause and changed their pitch in the quote. This is in line with the finding that children are usually able to use intonation functionally by the age of five (Wells, Peppé, & Goulondris, 2004).

In comprehension, children find it more difficult to use prosody as a cue to distinguish direct from indirect speech. Hewlett et al. (2003) studied the interpretation of prosodic cues in five- to twelve-year-old English-speaking children and a control group of adults. The participants listened to syntactically ambiguous speech reports like (30a) and (30b) that can only be disambiguated by their prosodic characteristics.

- (30) a. Pikachu said Squirtle was chasing the cat.
 b. "Pikachu", said Squirtle, "Was sleeping in the classroom".

The participants' task was to decide which character was performing the action (Squirtle in the indirect report of (30a), Pikachu in the direct report of (30b)). As typical for prosody comprehension (e.g., Wells, Peppé, & Goulondris, 2004), there was a lot of individual variation in the ability to interpret prosodic cues in all age groups, including the adults. The results show that the indirect speech interpretation was the default for children. Even twelve-year-old children were not adult-like in recognizing the acoustic cues for direct speech (Hewlett et al., 2003).¹¹

2.4.8 Written quotation marking

In writing, the context shift in direct speech can be signaled in several ways, most commonly by putting quotation marks around the quoted utterance. Several studies document that young writers have difficulties in marking quoted speech correctly (Bredel, 2004; Cordeiro, Giacobbe, & Cazden, 1983; Ferreiro & Pontecorvo, 1999; Ferreiro & Zucchermaglio, 1996). Initially, children do not separate quotations from the rest of the text at all. Later, they often use incorrect ways to signal quotations. A frequent error concerns the closing quotation marks, which are either omitted or incorrectly placed, like in (31) (Bredel, 2004; Cordeiro et al., 1983).

- (31) We are at my Aunt Susie's house. I said "Hello to Danny Stephen Mickey and Tommy." (Cordeiro et al., 1983, p. 327–328)

The correct use of quotation marks requires the ability to distinguish between direct and indirect speech. Even though in both report types another person's utterance is presented, only quotations need to be typographically set apart from the rest of the text. The reason for this is that direct speech involves a deictic shift to the reported speaker's perspective, which needs to be marked for the reader. This conceptual difference between direct and indirect speech seems to be difficult to grasp for children. We do not only find direct speech reports without any quotation marking, but also the opposite mistake that children overextend quotation marking to indirect speech reports, like in (32) (Bredel, 2004).

- (32) Ich habe gefragt: „Ob ich einen Freund mitnehmen kann“? „Ja aber nur einen!“
I asked: "if I can bring a friend"? "Yes but only one!" (Bredel, 2004, p. 231)

¹¹ Note that direct speech constructions like (30b) in which the reporting clause interrupts the quote are relatively rare in spoken language. This might be one of the reasons why the indirect speech interpretation was the default in this experimental setup.

Interestingly, such mistakes in written speech reports still occur at the age of eleven to twelve, which suggests a long developmental trajectory of quotation mark acquisition. This contrasts sharply with the prosodic marking of quotations, which is in place already at the age of three to four.

For children, face-to-face interactions are the primary form of communication and they allow some flexibility in switching between direct and indirect speech. The strict punctuation conventions in written language seem to be challenging for children because they presuppose a rigid direct–indirect distinction. Interestingly, quotation marks are a relatively new convention that was established in the transition from an oral to a literate culture (Maier, 2015b). I argue that similar to this historic development, children also need to accomplish a transition from orality to literacy, and with it a change from a flexible way of speech reporting to a strict direct–indirect distinction.

2.4.9 Distinguishing different discourse functions

So far, little is known about how different discourse contexts influence children's production and comprehension of speech reports. However, since researchers have studied children's speech reports in different communicative contexts such as narratives, playing with a doll house and spontaneous interactions with peers, we can compare preferences for direct or indirect speech in these discourse contexts.

In narratives, English-, Swedish- and Turkish-speaking children seem to produce more direct than indirect speech (Ely & McCabe, 1993; Nordqvist, 2000; Özyürek, 1996). This suggests that they choose the more vivid speech report type in order to engage their listeners in the story, similar to adults (Vincent & Perrin, 1999). However, the Danish children in Engberg-Pedersen and Christensen's (submitted) study do not fit this pattern: the ten- to thirteen-year-old typically developing children preferred indirect speech reports in their frog story narratives.

A possible explanation for this divergent finding could be that the Danish children in Engberg-Pedersen and Christensen's study are older than most children in the other studies. This explanation is supported by results from Nordqvist (2001a) who elicited and analysed frog story narratives from Swedish speakers of a wide age range (3–15, adults). She found that both younger children and adults preferred to report the characters' speech with direct speech reports. Older children above the age of nine, however, used mainly indirect speech, comparable to their Danish peers. Nordqvist (1998) offers the explanation that young adolescents use more indirect speech because they follow a *speak as you write* strategy: "information structuring that is typical of writing seems not only to have emerged, but is actually overgeneralized to the oral narratives" (Nordqvist, 1998, p. 44).

In addition to narrative contexts, children's use of speech reports has been studied in the context of pretend play activities. Nordqvist (2001a, 2001b) analysed the speech reports

that three-year-old Swedish children produced when playing with a doll house together with their mothers. She found that children clearly preferred to report the dolls' utterances with unframed direct speech reports, i.e., direct quotations without a reporting clause. Nordqvist explains the high number of reports without a reporting clause with the availability of non-verbal cues such as "holding the dolls, wiggling them, directing the gaze towards them", which render linguistic frames to identify the reported speaker unnecessary (Nordqvist, 2001b, p. 63). However, I argue that these unframed ascriptions of speech in the context of role play activities should not be classified as speech reports, but rather as role play utterances. In contrast to speech reports, a speaker in role play does not *report* a character's speech but pretends to actually *speak as* the character. Metalinguistic reporting frames are therefore not redundant, as Nordqvist suggests, but in fact counterproductive because they destroy the dramatic illusion that the doll is the one who is doing the talking. If we exclude all unframed direct speech reports from Nordqvist's dollhouse corpus, only 7 speech reports remain (6 direct and 1 indirect). This is not enough to derive meaningful conclusions (cf. Köder, to appear).

Boeg Thomsen (2014) studied the functions of speech reports that Danish children produced in a daycare setting. She found that children used direct speech reports mainly to create a dramatic effect in narratives, but they preferred indirect speech when regulating behaviour with authority, telling on others or accusing others of inconsistencies. Here is an indirect speech report from Boeg Thomsen's corpus that a three-year-old girl produced to appeal to the authority of her kindergarten teacher Sally:

- (33) Sally sagde at jeg skulle ikke sove.
'Sally said that I wasn't supposed to sleep' (3;3)

Boeg Thomsen's findings indicate that from early on children are aware of the functional differences of direct and indirect speech. This is a first step towards understanding children's preferences for direct or indirect speech in different discourse contexts. Because of the limited number of speech reports in Boeg Thomsen's corpus, further production studies based on bigger corpora are required to gain insights into children's pragmatic competence of using direct and indirect speech reports appropriately.

In chapter 3, I will provide such an analysis based on a large corpus of spontaneous Dutch and German child language. The amount of data and the diverse communicative situations in which the children interact with their caretakers provide the opportunity to systematically investigate the effect of different discourse contexts on children's use of direct and indirect speech. In addition to this production study, I will also experimentally test whether children's comprehension of speech reports is influenced by the discourse context (chapters 4, 5 and 6). As it will turn out, the discourse context plays an important role in both children's production and their comprehension of speech reports.

2.4.10 Children's difficulty with the direct–indirect distinction

In section 2.4, I have provided an overview of previous research on children's acquisition of speech reports. Previous studies indicate that children start to produce speech reports as early as two, but that the complete mastery of direct and indirect speech extends well into adolescence.

Children's difficulties in acquiring speech reports are not purely of a syntactic nature. The challenge children face seems to be related to distinguishing direct and indirect speech and the specific clusters of syntactic, semantic, prosodic, and pragmatic features associated with these two report types. In children's production of speech reports, this problem is reflected in the use of constructions that mix aspects of direct and indirect speech. Children combine, for instance, the syntax of indirect speech with the use of shifted pronouns, or make mistakes in the marking of quotations in written texts. In comprehension, children seem to allow reports that combine the extraction of *wh*-elements (indirect speech feature) with shifted indexicals (direct speech feature).

The review of previous studies revealed a clear bias towards production studies. Children's comprehension of direct and indirect speech has received little attention so far. Exceptions are studies on children's prosody perception in speech reports (Hewlett et al., 2003) and on *wh*-movement (Hollebrandse, 2007). In addition, research on children's comprehension of complement clauses (e.g., De Villiers & Pyers, 2002) provides insights into the interpretation of indirect speech reports, but not into the difference between direct and indirect speech understanding.

In this dissertation, I present three novel comprehension experiments that compare children's processing of pronouns in direct and indirect speech (chapters 4, 5 and 6). In view of earlier developmental studies, the question arises at what age children can distinguish direct from indirect speech and evaluate pronouns with respect to the correct speech context. My experimental results show that under certain pragmatic conditions, children evaluate pronouns in direct speech with respect to the reporting speech context, like in indirect speech. This direct-indirect mixing in comprehension provides support for the view that children operate with a more fluid concept of reported speech, not strictly distinguishing between direct and indirect speech.

2.5 PRONOUNS IN SPEECH REPORTS

The central question of this dissertation is how the semantic difference in perspective between direct and indirect speech affects children's and adults' processing of deictic expressions. I restrict my attention to pronouns for two reasons. First, in Dutch, pronouns in speech reports have a clear deictic orientation point: In direct speech they have to be evaluated with respect to the reported speaker's perspective and in indirect speech with

respect to the reporting speaker's perspective. By contrast, spatial and temporal indexicals and speaker adverbs are less strictly tied to one or the other speech context when they occur in speech reports (Harris & Potts, 2009; Plank, 1986). Second, children acquire pronouns before other context-dependent expressions such as spatial terms (*in back of/in front of*), demonstratives (*this/that*), locatives (*here/there*) or deictic verbs of motion (*come/go, bring/take*) (Tanz, 1980). This could be due to the fact that pronouns are more frequent than other deictic terms and that the different pronominal values have clearly defined boundaries (cf. *I vs. you* with *here vs. there*) (Tanz, 1980).

In the following sections, I first introduce some general features of pronouns (2.5) and subsequently highlight relevant findings on children's acquisition of pronouns (2.6).

2.5.1 Person distinctions

Deictic pronouns such as *I, you* and *he/she* are context-dependent expressions. This means that listeners need to have knowledge of the speech context – in particular the distribution of speech-act roles – in order to determine their meaning. The actual speaker and his or her time and place of utterance constitute the so-called *deictic center* or zero-point of the spatio-temporal coordinate system (Bühler, 1934; Lyons, 1977). However, because in a conversation speaker and addressee constantly switch roles, the deictic center also shifts and with it the meaning of deictic pronouns that are tied to the speaker's perspective.

We can distinguish between three main communicative roles in an interaction: speaker, addressee and other people. First-person *I* and second-person *you* refer to the primary participants of an interaction: speaker and addressee. Third-person *he* and *she* refer to a male or female person other than speaker and addressee (Lyons, 1977). Since both speaker and addressee are aware of their communicative roles, the referents of first-person and second-person pronouns are automatically salient in the discourse (Diessel, 2012). By contrast, in the case of third-person pronouns additional cues such as pointing gestures or a previous linguistic mention are necessary to establish unambiguous reference to an individual or object (Halliday & Hasan, 1976; Kaplan, 1989).

Based on these observations, Lyons (1977, p. 638f) claims that there is a “fundamental, and ineradicable difference between first-person and second-person pronouns, on the one hand, and third-person pronouns on the other”. This split between first- and second- versus third-person pronouns is widely accepted in different areas of linguistics such as formal semantics, typology and language acquisition research (e.g., Heim, 1991; Kaplan, 1989; Legendre & Smolensky, 2012; Schlenker, 2003; Siewierska, 2004). In Kaplan's (1989) framework, for instance, first- and second-person pronouns are pure indexicals, directly getting their reference from the context parameters. Deictic third-person pronouns, however, are genuine demonstratives, which need a demonstration (e.g., pointing) to be interpretable.

2.5.2 Deictic vs. anaphoric pronouns

The pronoun system is not only divided into first- and second-person versus third-person pronouns. Within third-person pronouns, linguists traditionally distinguish between deictic and anaphoric uses (Lyons, 1977). A pronoun is used deictically if it points to a person or thing in the extra-linguistic context; it is used anaphorically if it refers to a linguistic antecedent in the discourse (within or beyond the sentence level) to achieve reference.

As an example of a deictic third-person pronoun, consider the use of *she* in (34). The speaker utters the sentence *She is a genius!*, while pointing to the referent of *she*, Anna.

- (34) She [pointing to Anna] is a genius!
 (35) I told Anna_i that she_i is a genius.
 (36) I heard Anna_i's talk yesterday. She_i is a genius!

In examples (35) and (36), *she* is used anaphorically. In (35), the third-person pronoun *she* is bound by the linguistic antecedent *Anna* in the matrix clause. In example (36), *she* is not syntactically bound, but rather pragmatically bound or bound in the wider linguistic discourse.

Imagine now the following situation. After listening to Anna's talk, a colleague turns to you and says:

- (37) She is a genius!

Your colleague does not point to Anna, so *she* is not a genuine demonstrative in Kaplan's sense. Nevertheless, it is clear in this situation that your colleague uses *she* to refer to Anna because she is the contextually most salient female person. Both linguistic and extra-linguistic factors can contribute to Anna's salience: You have both just listened to her talk, Anna might be in your shared perceptual field, and you might also have linguistically mentioned her in a previous conversation.

This example indicates that there is no clear-cut boundary between deictic and anaphoric uses of third-person pronouns (cf. Levinson 2004). According to Hunter (2014), all expressions that can be sensitive to the extra-linguistic context, can also be sensitive to the linguistic context and vice versa. To provide a unified mechanism of contextual dependence, Hunter extends the formalization of linguistic context-sensitivity to cover also cases of extra-linguistic context-sensitivity. This does not only bridge the gap between anaphoric and deictic uses of third-person pronouns, but also between third-person *he* and *she* and the person indexicals *I* and *you*.

Indexicals and demonstratives, like classical anaphoric pronouns, serve to find an entity already available in the context so that new information about that entity can be added to the linguistic context. (Hunter, 2014, p. 41)

Crucial for the interpretation of first-, second- and third-person pronouns is contextual salience: Personal pronouns refer to the most salient individual with certain properties (e.g., non-participant, female), regardless of whether that salience is caused by linguistic information non-linguistic information or a combination of both (Hunter, 2014; von Heusinger, 2002).

2.5.3 Quoted pronouns

In terms of reference, pronouns embedded in direct speech are exceptional. The first-person pronoun *I* in (38), for instance, does not refer to the actual speaker who produced the speech report.

(38) Anna said, “I am a genius!”

What is more, according to Kaplan (1989), *I* in (38) is not referential at all because words within quotation marks are semantically inert, i.e., do not have their standard meaning. However, even though the content of the quotation is irrelevant for the truth conditions of the sentence – (38) is true iff Anna uttered a certain string of words –, it can still be relevant for the discourse as a whole (Recanati, 2000). In this sense, pronouns like *I* can retain their standard meaning in a quotation. *I* still refers to the speaker, albeit not to the actual reporting speaker, but to the speaker of the reported speech context, Anna.

In a typical face-to-face interaction, a listener can single out the current speaker by audio-visual information (e.g., mouth movement, direction of sound, voice). However, in order to identify the speaker in the non-actual reported speech context, listeners have to rely mainly on linguistic cues (Halliday & Hasan, 1976). In (38), the reported speaker and referent of quoted *I*, Anna, is explicitly mentioned in the reporting clause. This indicates that for the interpretation of deictic expressions embedded in direct speech, information available in the linguistic discourse is crucial.

In this dissertation, I study the processing of first-, second- and third-person singular pronouns in direct and indirect reported speech. I consider the effects of both linguistic and extra-linguistic factors on pronoun processing in speech reports. As we shall see, children’s and adults’ ease and accuracy of pronoun interpretation in speech reports depends on the type of discourse context, the perspective of the listener, and the presence/absence of a co-referential term.

2.5.4 Pronouns in Dutch and German

Personal pronouns can exhibit different kinds of features: person (e.g., speaker, addressee), number (e.g., singular, plural), gender (e.g., feminine, masculine), social rank/relationship (e.g., to express different degrees of formality, politeness, status) and case (e.g., nominative).

Since in the empirical studies I focus on Dutch¹² and German, I will briefly point out characteristics of the pronominal systems of these two languages. Dutch and German are typical in the respect that only third-person pronouns are gender-marked.¹³ In the second-person singular, German distinguishes between the informal form of address *du* and the formal form *Sie*. Similarly, Dutch exhibits a contrast between the informal *je/jij* and the formal *u*. Some pronouns in Dutch have a strong and a weak version, such as *zij* (strong) vs. *ze* (weak) in the third-person singular feminine. In the experimental studies, I will test children's and adults' comprehension of Dutch first-person (*ik* 'I'), second-person (*jij* 'you') and third-person (*hij* 'he', *zij* 'she') singular pronouns in subject position. Only the informal, strong forms of pronouns are used.

2.6 CHILDREN'S ACQUISITION OF DEICTIC PRONOUNS

When acquiring pronouns, children face several challenges that are related to the context-dependent meaning of these linguistic expressions. In the following sections, I identify relevant cognitive prerequisites for the acquisition of pronouns (2.6.1). Next, I summarize previous findings on how children learn to produce and comprehend personal pronouns (2.6.2, 2.6.3), before pointing out challenges specific to the use of pronouns in reported speech (2.6.4).

2.6.1 Prerequisites of pronoun acquisition

In order to understand the meaning of personal pronouns, children need to know that pronouns refer to people via the communicative role that they occupy at the time of utterance. Stawarska (2009) claims that the linguistically coded speaker and addressee roles have precursors in early nonverbal exchanges between children and their caretakers that take place via vocalizations and eye gaze. These so-called proto-conversations exhibit a turn-taking structure similar to verbal conversations (e.g., Jasnow & Feldstein, 1986; Trevarthen, 1979).

¹² I focus on Dutch as used in the Netherlands and not, for instance, in Flanders. While in the Netherlands, *je/jij* is used in the second-person singular, in Flanders *ge/gij* is preferred.

¹³ Of the 133 languages with gender-marked pronouns in Siewierska's (2004) sample, 97% have gender in the third person, 18% in the second person and 3% in the first person. The explanation for this finding is that gender-marking is redundant for interpreting first- and second-person pronouns because the referents of these pronouns are already sufficiently determined by the communicative role. However, for third-person pronouns the gender feature can be an important cue to reduce the set of potential referents (Siewierska, 2004).

It may be that the protospeaker and protoaddressee roles that take shape in synchronized face-to-face interactions in infancy serve as necessary (albeit by no means sufficient) preconditions of acquiring interpersonal deixis. (Stawarska, 2009, p. 112)

A second prerequisite not only for the acquisition of pronouns but for symbolic communication in general is the ability to participate in joint attentional scenes (Tomasello, 1999). This ability emerges at nine to twelve months alongside with the understanding of other persons as intentional agents (Tomasello, 1999). Joint attentional scenes consist of three elements: the child, the adult and an object of joint attention. Crucially, it is not sufficient that child and adult both attend to the same object; they also both have to be aware of the other person's attention to the object in question. Joint attentional scenes build the socio-cognitive foundations for scenes of reference. The adult can, for instance, use a word like *diaper* to direct the child's attention to an object in the shared perceptual space. Since the roles in the interaction are interchangeable, the child can use the same word to influence the adult's attentional states in turn (Tomasello, 1999).

However, while children can simply imitate the adults' use of words like *diaper*, the same strategy leads to systematic mistakes in the case of the deictic pronouns *I* and *you*. For instance, when a child wants to direct the mother's attention to her wet diaper by saying *Mommy, you peed in your diaper*, she incorrectly uses the second-person pronouns *you* and *your* to refer to herself. Pronoun reversal errors like this are well-documented in children's acquisition of personal pronouns in both spoken and sign languages (Chiat, 1982; Dale & Crain-Thoreson, 1993; Evans & Demuth, 2012; Oshima-Takane, 2009; Petitto, 1987). However, not all children go through a phase of pronoun reversing, and those children who do usually do not reverse pronouns consistently (Chiat, 1981, 1986; Dale & Crain-Thoreson, 1993).

An explanation why some children show pronoun reversal errors could be that they have not fully developed the perspective-taking skills necessary to perform a deictic shift. This is supported by two kinds of empirical evidence. First, the studies by Loveland (1984) and Ricard, Girouard, and Décarie (1999) indicate that perspective-taking skills are a prerequisite for the acquisition of deictic pronouns: the understanding of two spatial perspectives preceded the full mastery of first- and second-person pronouns, while the understanding of three spatial perspectives was prior or simultaneous to the acquisition of third-person pronouns. Second, autistic children have deficits in both pronoun acquisition (Evans & Demuth, 2012; Hobson, Lee, & Hobson, 2010; Kim, Paul, Tager-Flusberg, & Lord, 2014; Lee, Hobson, & Chiat, 1994) and cognitive perspective-taking (Baron-Cohen, Leslie, & Frith, 1985), which could indicate an underlying problem with shifting perspectives.

Several studies indicate that it helps children to figure out the context-dependent meaning of *I* and *you* when they have systematic opportunities to overhear conversations between other people from an external perspective (Oshima-Takane, 1988; Oshima-Takane, Takane, & Shultz, 1999). This is especially relevant for the acquisition of second-person pronouns:

Since each of the child's conversational partners uses first person pronouns in self-reference, the child is frequently exposed to the shifting reference of first person forms. In contrast, unless the child hears and attends to overheard speech, second person pronouns always seem to refer to the child, so there is less evidence for shifting reference. (Dale & Crain-Thoreson, 1993, pp. 575–576)

The fact that second-born children regularly witness conversations between their parents and their older sibling explains why they use deictic pronouns earlier than first-born children (Oshima-Takane, Goodz, & Derevensky, 1996). Autistic children's failure to attend to speech directed at someone else could be an alternative explanation for their frequent pronominal errors (Oshima-Takane & Benaroya, 1989).

2.6.2 Children's production and comprehension of pronouns

Let us now have a closer look at children's acquisition of first-, second-, and third-person pronouns. It is important to distinguish between the ability to produce *I*, *you* and *he/she* and the ability to correctly interpret these pronouns. In child language, different kinds of asymmetries have been observed: comprehension precedes production for many types of referring expressions (e.g., anaphoric subject pronouns in Dutch). But also the opposite pattern is attested, i.e., that production precedes comprehension (e.g., anaphoric object pronouns in Dutch) (Hendriks & Spenader, 2006; Hendriks, van Rijn, & Valkenier, 2007; Hendriks, 2014; Koster, Hoeks, & Hendriks, 2011). In the following, I will therefore first present previous studies on children's production of pronouns and after that turn to children's interpretation of pronouns.

Producing pronouns

Naturalistic observations and elicited production studies show that children typically start to produce first-, second-, and third-person singular pronouns at around the age of two (Bates, 1990; G. W. Bol & Kasparian, 2009; Bol & Kuiken, 1986; Charney, 1980; Imbens-Bailey & Pan, 1998; Pierce, 1992). There are individual differences in the routes that children take towards a full mastery of the pronominal system: Some children initially use names to refer to themselves and their addressees (e.g., *Lisa hungry*), thereby avoiding deictic shifts. Others produce pronouns from early on or switch between nominal and pronominal forms (Bates, 1990; Bloom et al., 1975; Smiley & Johnson, 2006).

A closer look at the order of acquisition reveals that children start with first-person pronouns, followed by second-person pronouns and anaphoric third-person pronouns, while deictic third-person pronouns emerge last (Charney, 1980). In terms of frequency, children show a clear preference for first-person pronouns (Bol & Kuiken, 1986; Imbens-Bailey & Pan, 1998; Nelson, 1975).

Interpreting pronouns

Comprehension studies in different languages such as English and French come to the same conclusion: Children understand first- and second-person pronouns before third-person pronouns (Brener, 1983; Charney, 1980; Legendre & Smolensky, 2012; Murphy, 1986). This is in line with the proposed theoretical distinction between first- and second-person pronouns versus third-person pronouns (Lyons, 1977). Third-person pronouns might be more difficult to interpret because their meaning is negatively defined as referring to neither speaker nor addressee (Lyons, 1977). Additionally, the person occupying the communicative role of non-participant might be perceptually less salient than speaker and addressee. As Brener (1983) points out, the speaker can be identified via the voice cue, the addressee via body orientation and facial expression, but the behavior of non-participants – if they are present at all – provides no clues about their communicate role.

It is crucial to note that the above-mentioned experiments tested children's comprehension of *deictic* third-person pronouns. There is evidence that *anaphoric* third-person pronouns are easier to understand for children.¹⁴ Not only is the anaphoric use common for third-person pronouns, production data also show that children themselves use *he* and *she* predominantly with an anaphoric function, i.e., to maintain the interlocutor's attention on a previously mentioned referent. Deictic uses to direct the listener's focus to a new referent are rare in child language and when they occur they are usually associated with a pointing gesture (Charney, 1980; Salazar Orvig et al., 2010).

In addition, Charney (1980) also found a difference in children's comprehension of first- and second-person pronouns. Children made significantly less mistakes in the interpretation of the second-person possessive pronoun *your* than of the first-person pronoun *my*, but only if they occupied the role of the addressee. As external observers, children's interpretation of first- and second-person pronouns did not differ (Brener, 1983; Charney, 1980). Taken together with the findings of her production study, Charney concludes that in both the role of the speaker and the role of the addressee, children acquire the pronoun first that

¹⁴ However, this is only true for pronouns in subject position. In several languages, as for instance English (Chien & Wexler, 1990) and Dutch (Koster, 1993; Spenader et al., 2009), children are delayed in the interpretation of object pronouns. Even after the age of six, children interpret the object pronoun *her* in a sentence like *Mama Bear is washing her* as referring to Mama Bear instead of to the other character, Goldilocks.

refers to themselves. This means that *I* precedes *you* in production, while *you* precedes *I* in comprehension. This is in line with Wechsler's (2010) *de se* theory of pronouns, that, in a nutshell, holds that self-reference is achieved via a different and supposedly easier mechanism than reference to other people.¹⁵

However, it remains rather mysterious that some children are able to produce first-person pronouns without also understanding them. An explanation for this puzzling finding is that pronouns might initially not be used as independent linguistic units, but as part of un-analyzed chunks such as *I-wanna*. This means that they are "used mechanically" rather than in their "fullest linguistic and grammatical meaning" (Merleau-Ponty, 1964, p. 151). This explanation is supported by Charney's empirical analysis.

2.6.3 Acquisition of pronominal gender marking

If we compare the acquisition of feminine and masculine third-person pronouns, we find that children tend to produce the masculine form *he* earlier and more frequently than the feminine form *she* in English, German and Dutch (Bol & Kuiken, 1986; Deutsch & Pechmann, 1978; Mills, 1986; Moore, 2001). The reason for this could be the higher number of *he* in children's input (Mills, 1986) and the fact that the masculine form also functions as gender-neutral, unmarked default form (Audring, 2009; Booij, 2002).

Comprehension studies suggest that gender-marking is a highly salient feature for children that guides their reference resolution from early on. Brener (1983) tested two- to five-year-old children's comprehension of singular pronouns in videotaped scenes including a speaker, an addressee and two non-participants (one male, one female). She found that children tended to choose a person with matching gender, even if they incorrectly selected the speaker or addressee as referent of third-person *he* or *she*. Arnold, Brown-Schmidt, and Trueswell's (2007) results indicate that English-speaking children can interpret gender-marked pronouns by the age of four. Additional eye-tracking data revealed that five-year-olds process gender information as rapidly as adults.

2.6.4 Challenges of pronoun interpretation in speech reports

In sum, typically developing children are able to produce and interpret first-, second- and third-person singular pronouns by the age of three to four. But what if pronouns are embedded in a direct or indirect speech report? The challenge then is that there are two possible deictic points of orientation: the reporting context and the reported context. Each of these speech contexts has its own speaker, addressee(s) and potentially other people. This double staffing of communicative roles could potentially confuse children and lead to direct-indirect mixes in both production and comprehension.

¹⁵ For a more detailed discussion of Wechsler's *de se* account, see section 5.4.2.

In production, children might produce reports that include features of both direct and indirect speech. In comprehension, the competition of speech contexts could lead to a systematic misinterpretation of pronouns. For instance, when hearing a report such as *Anna said, "I am a genius!"*, children might interpret *I* incorrectly as referring to the actual reporting speaker. Also the opposite mistake is possible. When listening to the report *Anna said that I am a genius*, children might assume that *I* refers to the reported speaker Anna. From a theoretical point of view, the interpretation of pronouns appears to be more difficult in direct than in indirect speech. While pronouns in indirect speech have to be interpreted with respect to the actual speech context – similar to simple non-embedded sentences –, pronouns in direct speech are anchored in the reported speech context. This means that listeners need to shift from the reporting speaker's perspective to the reported speaker's perspective. This perspective shift could be cognitively demanding for younger children.

So far, we have only empirical evidence that children mix features of direct and indirect speech in production (see section 2.4.4). It is unclear whether systematic direct–indirect mixes also occur in children's interpretation of pronouns in speech reports.

2.7 RESEARCH QUESTIONS AND OVERVIEW OF EMPIRICAL STUDIES

The overview of previous studies on speech reports and pronouns has revealed a research gap. The main semantic difference between direct and indirect speech, i.e., the difference in perspective, has not been investigated yet with psycholinguistic methods. Recent psycholinguistic studies tend to focus on the vividness distinction between direct and indirect speech (Stites et al., 2013; Yao et al., 2011, 2012; Yao & Scheepers, 2011). However, as pointed out in section 2.2.7, vividness is a subjective and gradual feature. In terms of perspective, however, direct and indirect speech are polar opposites: Pronouns in direct speech have to be evaluated with respect to the reported speech context (with the reported speaker as deictic center), whereas pronouns in indirect speech have to be evaluated with respect to the actual reporting speech context (with the reporting speaker as deictic center). Unlike languages such as Kwaza or Danish Sign Language, Dutch does not allow pronouns to deviate from this canonical ideal. Because of its clear-cut and syntactically well-marked direct–indirect distinction, I chose Dutch as language of investigation. This allows us to follow children from a more fluid picture with supposedly many direct–indirect mixes to the relatively rigid direct–indirect distinction in the adult grammar.

In the empirical studies, I address three main questions. First, do adults and children process pronouns in direct speech differently than in indirect speech? Second, which linguistic and non-linguistic factors influence the processing of pronouns in speech reports? And third, do children mix up direct and indirect speech in production and comprehension?

The empirical part consists of four studies: one corpus study and three psycholinguistic experiments. It covers both comprehension and production studies. This way, adults' and children's reported speech abilities can be examined from different perspectives and potential asymmetries between production and comprehension can be uncovered.

The purpose of the **corpus study** (chapter 3) is to find out when Dutch and German children start to produce different kinds of reporting constructions in natural interactions. I analyze spontaneous speech reports with the communication verb *say* in children between 1;1 and 4;6 years of age. By distinguishing direct speech reports of different degrees of complexity (clausal vs. non-clausal), a fairer comparison with indirect speech reports is possible. In German, it is particularly interesting to examine whether children acquire indirect speech with verb-final or verb-second word order first. I also compare the reports that children produce to those that they hear from their parents, siblings and other people. In addition, I present three pilot studies on specific aspects of speech reporting: (i) children's use of pronouns in direct and indirect speech, (ii) whose speech children report, and (iii) children's preferences for producing direct or indirect speech in different discourse contexts.

In **experiment 1** (chapter 4), I investigate adults' and children's processing of pronouns in direct and indirect speech. For that purpose, I have developed a referent-selection task, in which participants watch short animated scenes on a tablet. In each scene, one of three animals utters a direct speech report (e.g., *Elephant said, "I get the football"*) or an indirect speech report (e.g., *Elephant said that I get the football*) with a first-, second- or third-person pronoun. Participants have to select the referent of the pronoun by touching the respective animal on the screen. The central research question is whether pronoun interpretation is cognitively less demanding in direct or indirect speech. I first establish whether adults show a difference in accuracy and decision times between the interpretation of direct and indirect speech reports. After that, I follow children's development from 4 to 12 years.

In **experiment 2** (chapter 5), the experimental setup is similar to that of experiment 1. However, this time the test sentences are not presented as part of an interactive tablet game, but uttered by two hand puppets. The participant is directly involved in the interaction and not just an external observer as in experiment 1. The question is whether this alteration improves five- and nine-year-old children's interpretation of pronouns in speech reports. In addition, experiment 2 also contains a production part, in which the participants occupy the role of the reporting speaker. This allows us to compare the production and comprehension of speech reports in the same communicative setting. Will speakers' preferred report type in production also be the one that is easier to process for the listener?

In **experiment 3** (chapter 6), I investigate adults' and children's interpretation of pronouns in speech reports in the context of a narrative. Previous studies indicate that both adults and children prefer to produce the more vivid direct speech reports in narratives. But is it also easier for them to interpret pronouns from the character's perspective (direct speech) rather than from the perspective of the narrator (indirect speech)? In experiment 3, we simulated a picture book experience on a tablet. Four- and six-year-old children and adults listen to a story illustrated with pictures and have to interpret pronouns in direct and indirect speech reports.

In the **conclusion** (chapter 7), I connect the results from the four empirical studies to sketch a comprehensive picture of Dutch children's acquisition of the direct-indirect

distinction. My empirical findings indicate that even eleven-year-old children systematically misinterpret pronouns in direct speech under certain conditions. I identify three main factors that influence children's processing of pronouns in speech reports: the discourse context (information exchange vs. narrative), the communicative role of the listener (eavesdropper vs. addressee) and the presence or absence of a co-referential term in the reporting clause. By comparing the results from the comprehension and production studies, I discuss whether children's correct production precedes their correct comprehension of pronouns in speech reports. I end with methodological remarks and suggestions for promising future studies.

EMPIRICAL STUDIES

3

CORPUS STUDY: CHILDREN'S SPONTANEOUS PRODUCTION OF SPEECH REPORTS¹

The corpus study gives an overview of Dutch and German children's spontaneous use of speech reports. Knowledge about children's speech reporting in natural interactions is essential for the interpretation of the experimental results. It provides insights into children's acquisition of the direct-indirect distinction in everyday contexts, which are more diverse than the experimental situations of experiments 1, 2 and 3.

Since I test Dutch children's and adults' comprehension and production of speech reports in the experimental studies, it is evident why I look at Dutch children's natural use of reports in this chapter. In addition, I examine German children's production of speech reports because the acquisition of direct and indirect speech in German has not been investigated yet and the German child language corpus is more comprehensive than the Dutch one, which allows for in-depth quantitative analyses.

The chapter is divided into two parts. Part 1 (3.1) gives a general overview of Dutch and German children's production of speech reports in natural interactions. It examines when children start to use direct and indirect speech reports and how frequently they use both forms. In part 2 (3.2), I zoom in on a specific subclass of reports, namely reports in which a pronoun occupies the subject position of the reported sentence. These kinds of reports are suited best for comparison with the test sentences used in our comprehension experiments. I investigate what kind of pronouns children use in subject position (3.2.2), whose speech they report (3.2.3), and whether their preference for direct or indirect speech depends on the communicative situation (3.2.4).

3.1 DEVELOPMENTAL TRAJECTORY AND FREQUENCY

3.1.1 Introduction

So far, it is not clear at what age Dutch and German children acquire direct and indirect speech and how frequently they use the different reporting constructions in natural interactions. Chapter 3 will provide this fundamental information which will subsequently serve as the basis of reference for the experimental studies.

¹ Parts of chapter 3 are based on Köder (2013).

Previous studies on English, Swedish and Turkish suggest that children acquire direct speech before indirect speech (Ely & McCabe, 1993; Nordqvist, 2001a; Özyürek, 1996) and that they produce many mixed reports, in which they combine features of direct and indirect speech (Goodell & Sachs, 1992; Hickmann, 1993). Can similar observations be made for Dutch and German?

Both Dutch and German distinguish between direct and indirect speech reports. These closely related languages have similar indirect speech constructions with verb-final word order and an obligatory complementizer. However, the German language has an additional indirect speech construction with verb-second word order and no complementizer. So, in German the question is whether children produce indirect speech with verb-final or verb-second word order first. The results from previous studies are inconclusive. While Rothweiler (1993) claims that children produce verb-final constructions earlier, Brandt et al. (2010) argue that the order of acquisition depends on the specific complement-taking verb. However, a fatal flaw of their study is that they collapse all speech reports with verb-second word order. As a result, verb-second indirect speech becomes indistinguishable from direct speech reports of assertions. By drawing the relevant distinctions, I will determine the age of onset and the frequency of verb-final and verb-second indirect speech in German.

To round up the developmental analysis, I examine the speech reports of parents and other people when interacting with children, i.e., children's linguistic input. It will allow us to estimate to what extent children's use of reports is adult-like.

3.1.2 Data and Methods

The Dutch and German data is taken from the CHILDES database (MacWhinney, 2000). It encompasses verbal interactions between typically developing children and their interlocutors during everyday activities such as playing, book reading and lunch. The target children are 31 German-speaking² and 38 Dutch-speaking³ children between the ages of 1;1 and 4;6. There is a lot of variation in how often an individual child is recorded, at what exact age and how long the intervals between the recordings are.

For presentation purposes, the children are divided into seven age groups, covering a six months age range. Tables 1 and 2 show how many children are part of each age group and how many words (tokens) and utterances they produced. In both the Dutch and the German corpus, the densest data is available for children who are between 2;1 and 3;0 years old. This is precisely the time frame in which we expect to find children's first speech reports. The German child language corpus is 2.8 times larger than the Dutch one.

² The German child language corpus consists of the following CHILDES corpora: Caroline, Kerstin, Leo, Rigol, Simone, Szagun, Wagner.

³ The Dutch child language corpus consists of the following CHILDES corpora: BolKuiken, DeHouwer, Gillis, Groningen, Schaerlaekens, VanKampen.

The input is defined as all speech produced in the presence of the target children. Children's interlocutors are mainly their parents, but also grandparents, siblings and other people. On average, children hear 2.5 times more words (tokens) than they produce themselves.

Table 1: Overview of the German child language corpus

Age group	Children	Words	Utterances
1;1–1;6	28	18,667	13,864
1;7–2;0	29	81,724	55,906
2;1–2;6	30	312,325	149,162
2;7–3;0	27	300,905	92,337
3;1–3;6	14	182,553	48,799
3;7–4;0	11	92,939	24,101
4;1–4;6	5	60,618	16,254
Total	31	1,049,731	400,423

Table 2: Overview of the Dutch child language corpus

Age group	Children	Words	Utterances
1;1–1;6	2	417	389
1;7–2;0	25	51,866	31,447
2;1–2;6	22	95,564	42,589
2;7–3;0	23	110,755	39,498
3;1–3;6	18	78,878	26,288
3;7–4;0	4	17,118	5,072
4;1–4;6	2	14,034	3,601
Total	38	368,632	148,884

Previous studies have shown that children's development of complement clauses depends on the complement-taking verb, i.e., it is item-based rather than licensed by a general complement clause schema or rule (Brandt et al., 2010; Diessel & Tomasello, 2001; Diessel, 2004). This means that children might acquire different reporting verbs such as *say* or *ask* at different times and may have different preferences for using them with a direct or indirect speech report. Since our three experimental studies (to be discussed in chapters 4, 5 and 6) focus on children's and adults' comprehension of speech reports with the reporting verb *say*,

we also restrict the analysis of spontaneous speech to reports introduced by this reporting verb. This allows us to directly compare our findings from the comprehension experiments to the results of the corpus study.⁴

By means of the CLAN-program (MacWhinney, 2000), all occurrences of the lemmas *zeggen* 'say' (Dutch) and *sagen* 'say' (German) were identified in the Dutch and German corpus. In a next step, I eliminated manually all cases in which these verbs were not used as part of a speech report. These include, for instance, formulaic expressions such as *zeg* (*maar*) or *sag* (*mal*) 'say' in their function as attention getters.

In total, German children produced 702 speech reports with the reporting verb *sagen*, their interlocutors 1713 reports. Dutch children produced 247 speech reports with *zeggen*, their communication partners 1261.

I assigned each speech report to one of the six categories presented in table 3. I distinguish between different kinds of direct and indirect speech constructions. In the results section below, I explain the different categories and illustrate them with examples from the corpus. For the classification, I relied on word order and the presence or absence of a complementizer. I also took the linguistic context into account in order to get clues about the deictic orientation point of indexicals. Since no supporting audio or video data was available, I had no access to prosodic information.

In some cases, like for instance example (1), the linguistic features were not sufficient to determine whether a report is a form of clausal direct speech („*Die rosa Glasur ist kitschig*“, *hat Papa gesagt*) or V2 indirect speech (*Die rosa Glasur ist kitschig, hat Papa gesagt*). These reports were classified as 'ambiguous'.

- (1) Die rosa Glasur ist kitschig, hat Papa gesagt. (2;8)
'The pink icing is kitschy, Daddy said.'

⁴ *Say* has some other advantages: First, it is a generic communication verb and can be combined with both direct and indirect speech reports. Second, previous studies indicate that it is the most frequent communication verb among children (Ely & McCabe, 1993; Goodell & Sachs, 1992; Nordqvist, 2001a), which is useful for quantitative analyses.

Table 3: Classification of speech reports

Report type	Language	Example
Direct non-clausal	German Dutch	Anna sagte: "Juhu!". Anna zei: "Joehoe!". 'Anna said: "Yippee!".
Direct clausal	German Dutch	Anna sagte: "Ich bin glücklich". Anna zei: "Ik ben blij". 'Anna said: "I am happy".'
Indirect V-final	German Dutch	Anna sagte, dass sie glücklich ist. Anna zei dat ze blij is. 'Anna said that she is happy'.
Indirect V2		
a. Dependent main clause	German Dutch	Anna sagte, sie ist glücklich. ?Anna zei, ze is blij. 'Anna said, she is happy'.
b. Parenthetical frame	German Dutch	Sie ist glücklich, sagte Anna. Ze is blij, zei Anna. 'She is happy, said Anna'.
Ambiguous	German Dutch	Anna sagte es regnet. Anna zei het regent. 'Anna said it is raining'.
Mix (direct–indirect)	German Dutch	Anna ₁ sagte, dass ich ₁ glücklich bin. Anna ₁ zei dat ik ₁ blij ben. 'Anna ₁ said that I ₁ am happy'.

3.1.3 Children's first speech reports

Previous developmental studies have found that children produce direct speech before indirect speech (Ely & McCabe, 1993; Nordqvist, 2001a; Özyürek, 1996). However, this is hardly surprising considering that all kinds of linguistic and non-linguistic elements can be quoted, whereas indirect speech reports require a syntactic integration of the reported utterance, for instance in the form of a subordinate clause. In order to make the comparison between direct and indirect speech reports more informative, I distinguish two kinds of direct speech reports in my analysis: non-clausal direct speech and clausal direct speech. Non-clausal direct speech means that only linguistic and non-linguistic elements below the sentence level are quoted (e.g., *Anna said: "Yippee!"*). Clausal direct speech, by contrast, contains a quotation of a sentence with a finite verb (e.g., *Anna said: "I am happy"*). So, similar to indirect speech reports, clausal direct speech consists of two clauses: the reporting clause and the reported utterance. The interesting question is therefore not whether children acquire direct before indirect speech, but whether they acquire clausal direct speech before indirect speech.

The first speech reports that German and Dutch children produce with the reporting verbs *sagen* and *zeggen* 'say' occur around the age of two, which is in line with previous find-

ings for English and Swedish (Ely & McCabe, 1993; Nordqvist, 2001a). As expected, these reports are forms of non-clausal direct speech, i.e., direct speech reports of the syntactically simpler kind. Here are some of the earliest examples:

- (2) Thorsten sagt nicht „Hilfe“. (2;1) (German)
‘Thorsten doesn’t say “Help”.’
- (3) „Muh“ sagen die, ne? (2;1) (German)
‘“Moo” do they say, right?’
- (4) Klokken zeggen “tik tak”. (2;1) (Dutch)
‘Clocks say “tick tack”.’
- (5) Die suit [zegt] “boem”. (2;0) (Dutch)
‘That one says “boom”.’

At the age of two, children typically quote single words and onomatopoeic sounds like animal sounds (see (3)), sounds of objects (see (4)) and events (see (5)). The most frequently quoted words in the German corpus are *hallo* ‘hello’, *tshüss* ‘bye’, *danke* ‘thanks’, *bitte* ‘please’, *ja* ‘yes’, *nein* ‘no’ and *aua* ‘ouch’. Interestingly, when combined with *sagen* ‘say’, many of these words can function as surrogates for verbs children have not acquired yet. To give an example, *danke sagen* ‘say thanks’ in German can be used to replace *sich bedanken* ‘to thank’. This kind of speech reporting strategy offers children the opportunity to make use of their limited lexical resources in a communicatively efficient way.

Several months after their first non-clausal direct speech reports, children start to quote clauses with a finite verb. The first instance of clausal direct speech in our corpus can be found at the age of 2;5 in German (see (6)) and 2;10 in Dutch (see (7)).

- (6) „Will aber kleckern“, sagt der Joghurt. (2;5) (German)
‘“But want to spill”, says the yogurt.’
- (7) “Kleine olifant, mag ik nog een koekje”, zegt ie. (2;10) (Dutch)
‘“Little elephant, can I have another cookie”, he says.’

In my classification of speech reports, I distinguish two types of indirect speech constructions: indirect speech with verb-final word order and complementizer (indirect V-final) and indirect speech with verb-second word order (indirect V2). Verb-second indirect speech can be further subdivided into (a) constructions in which the reporting clause precedes the report, the so-called embedded main clause construction (Auer, 1998; Reis, 1997), and (b) constructions that contain a parenthetical frame positioned at the end or in the middle of the sentence. While all these constructions are grammatical in German, standard Dutch does not allow embedded main clauses (Zwart, 1997).

In our corpus, indirect speech with verb-final word order emerges at the age of 2;7 in German (see (8)) and 3;1 in Dutch (see (9)).⁵

- (8) Mechthild sagt, dass du Pipi machen sollst. (2;7) (German)
 ‘Mechthild says that you should do a wee-wee.’
- (9) Maar als mijn mama zegt dat jij niet aankomen mag. (3;1) (Dutch)
 ‘But when my mom says that you may not touch.’

Example (10) is the first instance of V2 indirect speech in the corpus, produced by a German girl at the age of 2;3.

- (10) Tiere sagen, ich bin leise. (2;3) (German)
 ‘Animals say I am quiet.’

While this report could syntactically also be an instance of clausal direct speech (*Tiere sagen: „Ich bin leise“*), the linguistic context suggests that it is in fact V2 indirect speech. Immediately after the girl has uttered (10), the mother repeats her utterance in indirect V-final form: *dass das schön ist, dass du leise bist* ‘that it is good that you are quiet’. This clarifies that the child is the one to whom the property of being quiet is ascribed. The first-person pronoun in (10) has hence an unshifted interpretation, which indicates indirect speech.

The examples presented above were the first instances of the different report types in the Dutch and German child language corpus. To get additional evidence for the order of acquisition, I followed the developmental path of an individual child. I selected the German-speaking child Leo whose linguistic development is exceptionally well documented due to frequent recordings and additional diary information. Leo produced his first instance of non-clausal direct speech at the age of 2;2.00. About 5 months later, at 2;7.01, Leo’s direct quotations contain a clause with a finite verb for the first time. Just two days later at 2;7.03, we find Leo’s first indirect speech report, a verb-final construction. At the age of 2;8.00, Leo also produced an indirect speech report with verb-second word order.

Based on the first occurrences in the corpus and Leo’s individual development, the following picture emerges. Dutch and German children start to produce speech reports with the reporting verbs *sagen* and *zeggen* ‘say’ from two years on. Their first speech reports are direct quotations of sounds or single words. Between the ages of two and three, the linguistic material children quote becomes more complex, in line with their evolving syntactic abilities.

⁵ It is worth noting that both clausal direct speech and V-final indirect speech occur 5 to 6 months earlier in the German than in the Dutch corpus. This time lag could be due to differences between the two data sets. The German child language corpus is not only larger, it also contains additional diary information on the boy Leo about his most complex utterances at the time. This increases the probability of detecting specific constructions earlier in German.

Our data show that the syntactically more complex speech reports, i.e., clausal direct speech, verb-second and verb-final indirect speech, emerge almost simultaneously in the second half of the third year and at the beginning of the fourth. This suggests that the order of acquisition is related to syntactic complexity. Direct quotations of linguistic elements below the sentence level (non-clausal direct speech) are syntactically easier than direct quotations of sentences or indirect speech reports. This is why children are able to produce them earlier.

3.1.4 Frequency of direct and indirect speech

3.1.4.1 General overview

After sketching the development of different reporting constructions, I now turn to the question how frequently Dutch and German children use these forms in natural interactions. Tables (4) and (5) give a quantitative overview of the different kinds of reports that children between the ages of 1;7 and 4;6 produced.

Table 4: Number (and percentage) of different speech reports per age group in German children

Age group	dir. non-clausal	dir. clausal	ind. V-final	ind. V2	ambig.	mix
1;7-2;0	0	0	0	0	0	0
2;1-2;6	56 (76.7%)	10 (13.7%)	0 (0%)	2 (2.7%)	5 (6.8%)	0 (0%)
2;7-3;0	163 (63.2%)	73 (28.3%)	7 (2.7%)	3 (1.2%)	12 (4.7%)	0 (0%)
3;1-3;6	95 (54.6%)	44 (25.3%)	7 (4.0%)	11 (6.3%)	17 (9.8%)	0 (0%)
3;7-4;0	60 (50.0%)	36 (30.0%)	4 (3.3%)	7 (5.8%)	13 (10.8%)	0 (0%)
4;1-4;6	46 (59.7%)	15 (19.5%)	2 (2.6%)	8 (10.4%)	5 (6.5%)	1 (1.3%)
Total	420 (59.8%)	178 (25.4%)	20 (2.8%)	31 (4.4%)	52 (7.4%)	1 (0.1%)

Table 5: Number (and percentage) of different speech reports per age group in Dutch children

Age group	dir. non-clausal	dir. clausal	ind. V-final	ind. V2	ambig.	mix
1;7-2;0	2 (100%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
2;1-2;6	59 (100%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
2;7-3;0	31 (81.6%)	6 (15.8%)	0 (0%)	0 (0%)	1 (2.6%)	0 (0%)
3;1-3;6	62 (72.9%)	17 (20.0%)	4 (4.7%)	1 (1.2%)	1 (1.2%)	0 (0%)
3;7-4;0	11 (44.0%)	12 (48.0%)	2 (8.0%)	0 (0%)	0 (0%)	0 (0%)
4;1-4;6	19 (50.0%)	18 (47.4%)	1 (2.6%)	0 (0%)	0 (0%)	0 (0%)
Total	184 (74.5%)	53 (21.5%)	7 (2.8%)	1 (0.4%)	2 (0.8%)	0 (0%)

The most striking result is that both German and Dutch children clearly prefer the use of direct speech. In German, 85% of the reports are (non-clausal and clausal) direct speech reports, in Dutch even 96%. The majority of the direct speech reports are of the simpler ‘direct non-clausal’ type and do not contain a finite verb. With respect to the frequency of indirect speech reports, we can observe a difference between Dutch and German children. While Dutch children prefer the indirect verb-final construction, German children produce more indirect verb-second constructions. Direct–indirect mixes are non-existent in Dutch and rare in German.

Since the sub-corpora of the six age groups differ in size, absolute frequencies cannot be compared across age groups. This is why I calculated a relative frequency measure that indicates the number of different speech report constructions per 10,000 utterances. The results are displayed in figures 1 and 2.

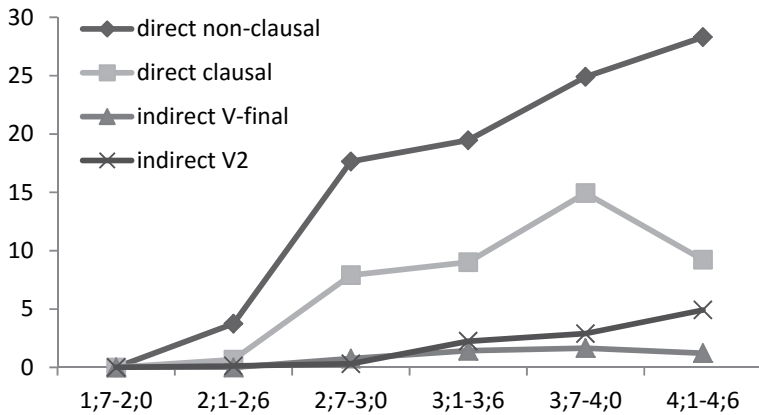


Figure 1: Relative number of speech reports with matrix verb *sagen* 'say' per 10,000 utterances in German children

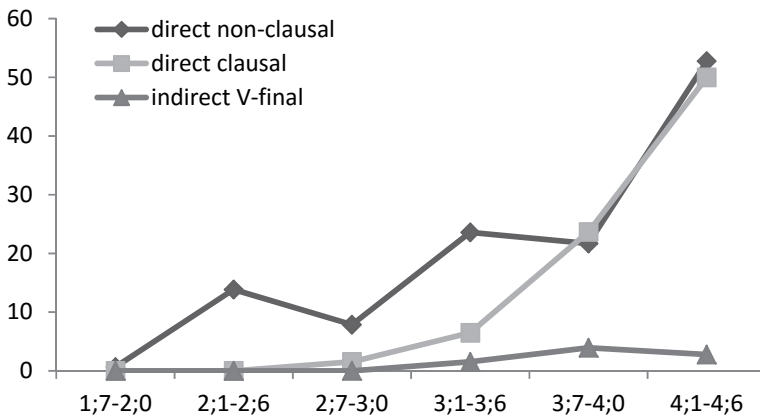


Figure 2: Relative number of speech reports with matrix verb *zeggen* 'say' per 10,000 utterances in Dutch children

The graphs indicate that the number of speech reports introduced by *sagen* and *zeggen* increases between the ages of 1;7 and 4;6. The most frequent report type is the syntactically simpler non-clausal direct speech. It dominates in German among all age groups and in Dutch until the age of 3;7, when clausal direct speech reaches comparable frequency levels. The number of indirect speech constructions stays at a relatively low level throughout children's early production of speech reports.

In the following sections, I discuss in detail the three main results of this quantitative overview: the predominance of direct speech (3.1.4.2), the differential preference for indirect V2 or indirect V-final in Dutch and German children (3.1.4.3), and the surprising shortage of direct–indirect mixes (3.1.4.4).

3.1.4.2 Direct speech dominates

As we have seen, German and Dutch children produce considerably more direct than indirect speech reports in natural interactions. This preference has also been observed in English- and Swedish-speaking children (Ely & McCabe, 1993; Nordqvist, 2000). Why do children use direct speech more often than indirect speech? I discuss two, not mutually exclusive explanations: (i) direct speech is easier to produce, and (ii) the discourse contexts in which children's reports are situated favor the use of the more vivid direct speech.

Easier in production. The first explanation for the higher frequency of direct speech reports is that the production of a direct speech report is cognitively less demanding than the production of an indirect speech report. The argument for this claim – originally presented by Li (1986) and adopted by child language researchers such as Goodell and Sachs (1992) and Nordqvist (2000) – is that direct speech only involves “reproducing” or “mimicking” a prior utterance. By contrast, indirect speech requires “rephrasing” or “paraphrasing” the speech of the reported speaker, which is associated with a higher processing effort. In my view, this line of reasoning is problematic because of its underlying assumptions: (i) that all speech reports are based on a previous utterance, and (ii) that children keep an utterance in mind and later reproduce it verbatim.

First, not all of children's speech reports are reports of past events. In fact, our corpus contains many reports that do not refer back to a specific real-world counterpart in the past. Children project, for instance, speech onto fictional characters (see (11)) or report what someone should say in the future (see (12)).

- (11) „Will aber kleckern“, sagt der Joghurt. (2;5) (German)
 “‘But want to spill”, says the yogurt.
 (12) jij moet zeggen: “ik zie een leeuw”. (4;4) (Dutch)
 ‘you must say: “I see a lion”

Without a particular model utterance in the past that the child has witnessed, the concept of “mimicking” or “reproducing” becomes obsolete (cf. Clark & Gerrig, 1990; Tannen, 2007; Wade & Clark, 1993).

Second, even if we assume that a previous utterance is referred to in the report, the idea of mimicking does not seem to comply with the psychological reality of human memory. Several studies show that people remember the gist of an utterance rather than its exact lin-

guistic form (e.g., Lehrer, 1989; Wade & Clark, 1993). In light of this research, it seems more appropriate to think of speech reports as (re)constructions rather than verbatim renditions or copies (Tannen, 2007).

Even if “mimicking” cannot account for the higher frequency of direct speech, the basic claim that direct speech is easier to produce than indirect speech could still be valid. In the following, I present perspectival and syntactic arguments to explain why direct speech might be less demanding for children than indirect speech.

Children might prefer direct speech because it allows them to take the perspective of the reported speaker. This claim might at first sound surprising, given that many studies document young children's difficulties in shifting from their own egocentric perspective to that of someone else – which seems exactly what is required in the production of direct speech reports. To give examples, children struggle with mental perspective-taking (Flavell, Flavell, & Green, 1983; Gopnik & Astington, 1988; Wimmer & Perner, 1983) and visual perspective-taking (Masangkay et al., 1974) until the age of four to five. However, as Moll and Meltzoff (2011) point out, it is crucial to distinguish the ability to take another person's perspective on an object from the ability to confront two divergent perspectives on the same object. While the theory-of-mind and visual perspective-taking tasks mentioned above require the confrontation of perspectives, I suggest that the production of direct speech reports requires only the shifting to someone else's perspective. This ability is well in place by the age of three (Moll & Meltzoff, 2011).

What is more, typically developing children even seem to enjoy taking other people's perspective, as evident from their early and frequent participation in role play activities. From two to three years on, they play the role of a superhero, pretend that their toy figures speak and converse with their imaginary friend – in case they have one (Harris, 2000; Taylor, 1999; Wolf, Rygh, & Altshuler, 1984). Enacting the (linguistic) behavior of someone else seems to be such a fundamental mechanism that children even fall back on the presumably easier enactment mode when telling a story (Hickmann, 1993; Nordqvist, 2001b). Children might prefer direct speech because of its similarities with utterances in role play. In both cases, a speaker shifts to the perspective of another person – a character in role play and the reported speaker in direct speech – and produces speech from that person's deictic point of orientation (Köder, to appear).

A second argument why children might prefer to produce direct speech reports is that in direct speech, specific aspects of the reported utterance can be demonstrated (Clark & Gerrig, 1990). To give an example, even with a limited lexicon children can imitate the stuttering of a person in a direct quotation. In indirect speech, they would have to use the descriptive term *stutter*, which they might not yet have acquired. Similarly, the illocutionary force of questions and commands can be demonstrated in direct speech, while in indirect speech difficult syntactic adjustments are required. Opting for direct speech therefore seems to be a successful strategy for children to maximize understanding.

Interestingly, this line of reasoning is supported by research on adult speakers with aphasia whose language is limited in lexical diversity and syntactic complexity. Groenewold, Bastiaanse, and Huiskes (2013) found that aphasic speakers use significantly more direct speech than non-brain-damaged adults. They suggest that speakers with aphasia use direct speech “to get around grammatical problems and word-finding difficulties” (Groenewold et al., 2013, p. 546).

Discourse context. The second explanation why children predominantly produce direct speech is that the discourse contexts in which these interactions are embedded favor the use of direct speech. Previous studies on both adults and children indicate that speakers do not just randomly use direct or indirect speech, but that they adapt their choice of report type to the communicative function of the report. Speakers tend to prefer direct speech to entertain or amuse their listeners, and indirect speech to inform their listeners about the propositional content of an utterance (Boeg Thomsen, 2014; Vincent & Perrin, 1999; Wade & Clark, 1993).

I hypothesize that the reason why children use more direct than indirect speech in our corpus is that they are mostly recorded during interactions such as story-telling and playing, in which the vividness of direct speech is a virtue. In the fictive world of narratives or play, the main purpose of a speech report is not to convey a certain proposition, but to (re)enact or simulate speech in an entertaining and amusing manner. This can be more easily achieved with a vivid direct speech report.

If these considerations are correct, the proportion of direct and indirect speech should depend on the discourse context. Based on the literature, fictive contexts like narratives or pretend play should yield a higher percentage of direct speech than more reality-oriented contexts, in which the transmission of information is central. In section 3.2.4, I examine the contextual embedding of children’s spontaneous speech reports more closely and put the discourse context hypothesis to the test.

3.1.4.3 Indirect speech with verb-second and verb-final word order

The overview of reported speech frequency in section 3.1.4.1 has revealed that German and Dutch children prefer different types of indirect speech constructions. Dutch children almost exclusively produce indirect speech reports with verb-final word order and complementizer. German children, however, use both V-final and V2 indirect speech. This difference can be explained by the fact that the indirect speech construction with an embedded main clause is not considered grammatical in standard Dutch (Zwart, 1997).

However, there is one exception in the Dutch child language data that shows many similarities with the German embedded verb-second construction:

- (13) Jij moet zeggen, ik ben zo mooi. (4;1) (Dutch)
 ‘You must say, I am so beautiful.’

Report (13) has verb-second word order, lacks a complementizer, and the response of the mother (*je bent ontzettend mooi* 'you are incredibly beautiful') indicates that the first-person pronoun refers to the reporting speaker. All these features suggest that (13) is an instance of V2 indirect speech. Alternatively, it could be a direct–indirect mix, combining the main clause syntax of direct speech with the deictic perspective of indirect speech. In order to assess which one of these options is more likely, it is important to know whether Dutch children's interlocutors use the embedded V2 construction in natural conversations. As we will see in section 3.1.5, there are no instances of embedded V2 constructions in children's input, so (13) is likely to be a direct–indirect mix.

Let us now focus on German children's indirect speech reports. In section 3.1.3, we have seen that German children start producing indirect speech with verb-second and verb-final word order at around the same age. However, in terms of frequency, German children prefer indirect speech reports with verb-second word order: 61% of German children's indirect reports with the reporting verb *sagen* 'say' have verb-second word order, 39% verb-final word order.⁶ The German indirect verb-second constructions can be further divided into embedded main clause constructions with a sentence-initial reporting clause, and verb-second constructions with a parenthetical reporting frame at the end or in the middle of the sentence. Of the 31 cases of V2 indirect speech in German, the majority (n=26) is of the embedded main clause type, such as (14).

- (14) Ich hab' dir ja gesagt, ich gewinne. (3;10) (German)
'I told you I'd win.'

Five reports have a parenthetical reporting frame, positioned either at the end or in the middle of the report, such as in example (15).

- (15) Die will ja, sagt se, die will auch mit Bilderbuch angucken. (3;1) (German)
'She wants, she says, she also wants to look at picture book.'

Taken together, this shows that German children prefer to use indirect speech reports with verb-second word order, more specifically embedded main clause constructions with a fronted matrix clause.

3.1.4.4 Not many direct–indirect mixes

Previous studies indicate that children have a less rigid direct–indirect distinction than adults. This is, for instance, apparent from their production of reports that lie in between direct and indirect speech (Goodell & Sachs, 1992; Hickmann, 1993). Against this background, it is quite

⁶ Note that the preference for the verb-second construction is probably even stronger considering that its frequency tends to be underestimated in our data. Some indirect V2 reports might have been classified as 'ambiguous' because they were indistinguishable from clausal direct speech reports.

surprising that the Dutch and German children in this study produced almost no cases of direct–indirect mixing. In addition to the Dutch report (13), (16) is the only candidate for a direct–indirect mix.

- (16) sagt der Peter eine Schokolade ich gegess habe (2;5) (German)
 ‘says Peter a chocolate I eaten have’

The German-speaking child who produced (16) combined the following features: verb-final word order (feature of indirect V-final), lack of a complementizer (feature of indirect V2, clausal direct), fronting of the direct object *eine Schokolade* ‘a chocolate’ (feature of indirect V2, clausal direct). Unfortunately, it is unclear from the context whether the first-person pronoun *ich* ‘I’ refers to the reported speaker Peter or the reporting speaker, the child. Only in the first case, (16) would qualify as a direct–indirect mix. Otherwise, it would be an ungrammatical indirect speech report that combines features of V-final and V2 indirect speech.

What factors are responsible for the diverging finding that our corpus contains so few direct–indirect mixes, whereas in Goodell and Sachs’s (1992) study approximately 22% of the four-year-old’s indirect reports were identified as mixes? I argue that this is due to two main differences between our studies: (i) a different definition of what counts as a direct–indirect mix, and (ii) a different corpus and methodology.

The first difference between our corpus study and Goodell and Sachs’s (1992) study is that we operate with different definitions of a direct–indirect mix. Goodell and Sachs seem to have a broad notion of mixing. They count, for instance, reports with backshifting mistakes, syntactic mistakes (e.g., concerning the indirect report of questions and commands), and lexical mistakes (e.g., wrong choice of speech act verb) as mixes. In my view, many of these ungrammatical reports are not a result of children’s inability to distinguish direct from indirect speech, but they arise from children’s difficulties with the syntax of complementation. According to my strict definition, a report is classified as a direct–indirect mix only if it combines syntactic features of one report type with the deictic features of the other report type. Applying this definition to Goodell and Sachs’s data reduces the number of mixes considerably, but there are still quite a few left.

The second difference concerns the data and methodology. I have analyzed reports that children produced spontaneously in natural interactions. In the studies of Goodell and Sachs (1992) and Hickmann (1993), however, children had to retell a conversation immediately after they listened to it. Children may be more prone to mixing, when parts of the exact wording and syntactic structure of the original are still fresh in their memory.

Consider, for instance, the details of Goodell and Sachs’s study, which exhibits the highest number of direct–indirect mixes. In one of their conditions, children had to retell a dialogue presented in the form of indirect speech reports. The exposure to indirect speech reports could have primed children to use more syntactic features of indirect speech in their

own narratives (e.g., the complementizer *that*) (cf. Serratrice, Hesketh, & Ashworth, 2015)⁷. Such a priming effect combined with children's general preference for direct speech could explain why in Goodell and Sachs's study, children often combined the syntax of indirect speech with shifted pronouns.

It is important to note that in studies on children's spontaneous speech reports there is no mention of direct–indirect mixes (Ely & McCabe, 1993; Nordqvist, 2001a). This confirms our suggestion that aspects of the elicitation procedure might be responsible for the high number of direct–indirect mixes in Goodell and Sachs's study. To provide further insights into children's production of speech reports, I will elicit and analyse speech reports from Dutch children and adults in experiment 2 (chapter 5). As we shall see, direct–indirect mixes are also rare in this experimental setting.

3.1.5 Input analysis: Reports of children's interaction partners

After having analyzed children's speech reports, it is time for a closer look at the reports of their interaction partners. In our corpus, children interact predominantly with their mothers, but also with their fathers, grandparents, siblings, peers or child language researchers. The linguistic input that children are exposed to helps us to answer the following questions: Do children's interlocutors also prefer direct speech? What functions do their speech reports have? Do Dutch adults produce embedded main clause constructions?

Tables 6 and 7 show the frequency of different report types with the reporting verbs *sagen* and *zeggen* 'say' in the German and Dutch input. The results are divided into seven groups corresponding to the age of the target children in whose presence the reports were uttered.

The most striking result is that children's interaction partners also produce significantly more direct than indirect speech reports. The German input consists of 84% direct speech, 10% indirect speech and 5% ambiguous reports. In Dutch, children's interlocutors produced 91% direct speech, 7% indirect speech and 1% ambiguous reports. The fact that adults and children prefer to use direct speech when interacting with each other is consistent with the discourse context hypothesis. Since the adult-child interactions are predominantly of a playful nature, speakers might opt for the more vivid and entertaining direct speech reports.

⁷ In their priming study, Serratrice et al. (2015) show that five-year-old English-speaking children who are exposed to indirect speech reports produce more indirect speech reports in their own narratives than children without a systematic exposure to indirect speech reports.

Table 6: Number (and percentage) of different speech reports in input to German children

Age of child listening to speech	Direct non-clausal	Direct clausal	Indirect V-final	Indirect V2	Ambiguous
1;1–1;6	174 (91.1%)	13 (6.8%)	1 (0.5%)	2 (1.0%)	1 (0.5%)
1;7–2;0	356 (63.6%)	156 (27.9%)	12 (2.1%)	13 (2.3%)	23 (4.1%)
2;1–2;6	110 (31.7%)	191 (55.0%)	12 (3.5%)	16 (3.5%)	18 (5.2%)
2;7–3;0	60 (23.9%)	116 (46.2%)	20 (8.0%)	40 (15.9%)	15 (6.0%)
3;1–3;6	56 (47.1%)	36 (30.3%)	5 (4.2%)	10 (8.4%)	12 (10.1%)
3;7–4;0	32 (21.6%)	76 (51.4%)	9 (6.1%)	16 (10.8%)	15 (10.1%)
4;1–4;6	37 (38.1%)	31 (32.0%)	15 (15.5%)	5 (5.2%)	9 (9.3%)
Total	825 (48.2%)	619 (36.1%)	74 (4.3%)	102 (6.0%)	93 (5.4%)

Table 7: Number (and percentage) of different speech reports in input to Dutch children

Age of child listening to speech	Direct non-clausal	Direct clausal	Indirect V-final	Indirect V2	Ambiguous
1;1–1;6	25 (92.6%)	2 (7.4%)	0 (0%)	0 (0%)	0 (0%)
1;7–2;0	277 (75.3%)	68 (18.5%)	22 (6.0%)	1 (0.3%)	0 (0%)
2;1–2;6	229 (55.3%)	150 (36.2%)	25 (6.0%)	3 (0.7%)	7 (1.7%)
2;7–3;0	103 (47.7%)	84 (38.9%)	21 (9.7%)	3 (1.4%)	5 (2.3%)
3;1–3;6	54 (31.4%)	105 (61.0%)	10 (5.8%)	0 (0%)	3 (1.7%)
3;7–4;0	16 (44.4%)	16 (44.4%)	2 (5.6%)	1 (2.8%)	1 (2.8%)
4;1–4;6	13 (46.4%)	9 (32.1%)	4 (14.3%)	1 (3.6%)	1 (3.6%)
Total	717 (56.9%)	434 (34.4%)	84 (6.7%)	9 (0.7%)	17 (1.3%)

However, it is important to point out that adults and children do not necessarily use direct speech reports for the same reasons. For instance, it could be the case that adults use direct speech to attract and maintain children's attention, whereas children use direct speech because it is easier for them. To find out more about the functions of speech reports that are directed towards children or uttered in their presence, I will now move from the quantitative overview to concrete examples.

In tables 6 and 7, we can see that German and Dutch children's conversation partners produce a lot of non-clausal direct speech when the target children are between 1;1 and 2;0 years of age. This is exactly the time frame before children produce the first speech reports themselves. What are the functions of adults' non-clausal direct speech reports to children of this age? Many non-clausal direct speech reports resemble example (17). Parents use these kinds of reports to connect language to children's perception.

- (17) “boem” zegt de bal (age of target child: 2;0) (Dutch)
 ‘“boom” says the ball’
- (18) sag mal „danke“ (age of target child: 1;4) (German)
 ‘say “thanks”’
- (19) zeg maar “ik heet Tomas” (age of target child: 1;8) (Dutch)
 ‘say “my name is Tomas”’

We can also find many direct speech reports of the form *sag (mal) x* or *zeg (maar) x* ‘say x’, where x can be a word (see (18)) or a phrase (see (19)). Parents typically use the ‘say x’ construction to prompt their children to produce a specific utterance that they deem appropriate in a certain communicative situation. The advantage of a direct speech report in these cases is that children are presented with a model utterance in the quotation that they can subsequently copy. Imagine that the direct speech report (19) *zeg maar “ik heet Tomas”* ‘say “my name is Tomas”’ would instead be formulated as an indirect report such as *zeg maar dat je Tomas heet* ‘Say that your name is Tomas’. In order to follow this request, the child would need to make deictic changes (from *je* ‘you’ to *ik* ‘I’) and syntactic changes (from verb-second to verb-final word order). Since children's pronominal system and syntax is not yet fully developed at the age of two, the risk of pronoun reversals and syntactic mistakes would be high.

In addition, parents frequently use direct speech reports to demonstrate what the child said (see (20)) or is able to say (see (21)).

- (20) [*mother to child*] omdat jij zei “mama, dragen, mama Thijsje dragen” (age of target child: 2;2) (Dutch)
 ‘because you said “Mommy, carry, Mommy carry Thijsje”’
- (21) [*mother to investigator*] ja, hij kan inderdaad nu “ik” zeggen (age of target child: 1;10) (Dutch)
 ‘yes, he can indeed say “I” now’

Again, speakers can accomplish this purpose better with a direct rather than an indirect speech report because not only the content is relevant, but aspects of the lexical and syntactic form of the child's utterance as well. For instance, while direct speech allows the adult speaker to present the morpho-syntactic errors in the child's utterance in example (20), the use of indirect speech would normalize the utterance and therefore obscure these specific features.

I now turn to the indirect speech reports that children's communication partners use. German adults prefer to use indirect speech with a verb-second construction (58%), similar to children. Again the majority of the indirect verb-second constructions (91 of 102) are embedded main clauses with a sentence-initial matrix clause.

In the Dutch input, the indirect verb-final construction is clearly the preferred indirect speech type ($n = 84$). However, there are also 9 indirect speech reports with verb-second word order. These indirect V2 reports are connected to a reporting clause positioned in the middle ($n = 1$) or at the end ($n = 8$) of the sentence, such as in example (22).

- (22) ze heeft nog niet gepoept, zegt ze
'she has not pooped yet, she says'

We did not find embedded main clauses with a fronted matrix clause in the speech of Dutch children's interlocutors. However, it is possible that some of the ambiguous reports are in fact embedded verb-second clauses.

3.2 ZOOMING IN ON THE PRONOUN CORPUS: THREE PILOT STUDIES

After this general overview of children's and adults' speech reports, I take a closer look at a specific subclass of these reports: children's direct and indirect speech reports with a pronoun in subject position of the reported utterance (e.g., *Anna said she is happy*). I zoom in on these kinds of reports because they resemble the test sentences that are used in the three experimental studies. This allows us to directly contrast children's production and comprehension of pronouns in speech reports.

I present three pilot studies in which I focus on specific aspects of children's speech reports: (i) what kinds of pronouns children use in direct and indirect speech (3.2.2), (ii) whose speech they report (3.2.3), and (iii) how the discourse context and function of the report influences children's choice of a direct or indirect speech report (3.2.4).

3.2.1 The pronoun corpus

I created a sub-corpus of the speech report corpus used in section 3.1. It consists only of children's speech reports that contain a pronoun as subject of the reported utterance, such as (23). In the following, I will refer to this corpus as the 'pronoun corpus'.

- (23) Die Martina hat gesagt, sie wollte in 'n Garten. (3;5) (German)
'Martina said she wanted (to go) into the garden.'

Similar to the general speech report corpus, all reports in the pronoun corpus contain the reporting verbs *sagen* or *zeggen* 'say'. In order to be able to compare the use of direct and indirect speech reports, I ensured that the same kinds of speech acts can be presented with both report types. Since in indirect speech, questions are usually reported with a different complement-taking verb (e.g., *fragen* (German) or *vragen* (Dutch) 'ask'), I excluded all direct quotations of questions.

In total, the pronoun corpus contains 116 (76 direct, 40 indirect) speech reports produced by German children and 34 (29 direct, 5 indirect) speech reports produced by Dutch children.⁸ In the following analysis, I focus on the German data because it contains a higher number of target sentences, which makes it more suitable for quantitative observations. I will indicate whether similar trends can be observed in the Dutch data.

3.2.2 Study 1: Children's use of pronouns in speech reports

In the three comprehension experiments discussed in chapters 4, 5 and 6, I examine children's interpretation of first-, second- and third-person singular pronouns in sentences such as *Hond zei: "Ik krijg de voetbal"* 'Dog said: "I get the football"'. Against this background, it is relevant to see what pronouns children use themselves when they are the reporting speaker.

Figure 3 gives an overview of how frequently German children use different kinds of pronouns in subject position of the reported utterance. It shows that German children produce more singular than plural pronouns. It is also apparent that children have different pronominal preferences in direct and indirect speech reports.

In direct speech, German children used mainly the first-person singular pronoun *ich* 'I' (54%), followed by the second-person singular pronoun *du* 'you' (26%). Noticeably, no third-person singular pronouns were found in direct speech. In indirect speech, however, third-person singular pronouns are the most frequently used pronouns (60%).⁹ It is important to draw attention to the fact that third-person pronouns in indirect speech are potentially ambiguous. In an indirect speech report such as *Anna said that she is happy*, third-

⁸ Note that reports containing elliptical pronouns are also included in the study. German children produced nine elliptical pronouns, all occurring in direct speech reports (e.g., „*will raus*“, *sagt das Nilpferd*“ ‘want out’, said the hippo’ (2;6)). With one exception, they were produced by children under the age of three.

⁹ In Dutch, the distribution is similar: Dutch children used also mainly first-person (59%) and second-person pronouns (41%) in direct speech, while avoiding third-person pronouns. In indirect speech, they also preferred the use of third-person pronouns.

person *she* can either refer to the subject of the matrix clause *Anna* or to another female person salient in the linguistic or extra-linguistic context. Of the 24 third-person pronouns in indirect speech, half is co-referential with the matrix clause subject.¹⁰

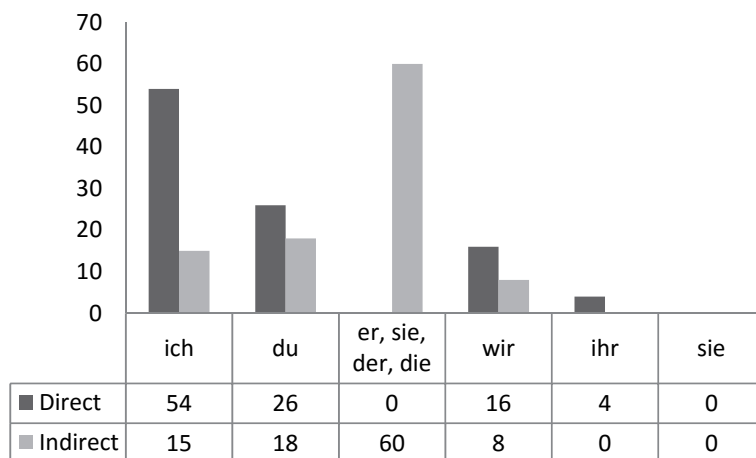


Figure 3: Percentage of different kinds of pronouns in subject position of the reported utterance in German children's direct and indirect speech reports.

Despite the observed differences between direct and indirect speech, the unifying element is that in both cases children often report the speech of a person who is talking about him- or herself. Because of perspectival differences between direct and indirect speech, reference to the reported speaker has to be established with different pronouns: a first-person pronoun in direct speech, and an anaphoric third-person pronoun in indirect speech.

¹⁰ German has two types of third-person pronouns: the personal pronouns *er, sie, es* and the demonstrative pronouns *der, die, das*. Of the 24 third-person pronouns in indirect speech, 15 are personal pronouns (9 feminine, 6 masculine), and 9 are demonstrative pronouns (2 feminine, 7 masculine). In German, personal pronouns prefer antecedents that are discourse topics (typically in subject position), while demonstrative pronouns have a strong tendency to avoid discourse topics (subject antecedents) as referents (Bosch, Katz, & Umbach, 2007; Bosch, Rozario, & Zhao, 2003; Bosch & Umbach, 2007; Comrie, 1997). The German children in the pronoun corpus do not seem to functionally distinguish the two types of pronouns. They used, for instance, also demonstrative pronouns to anaphorically refer back to the matrix clause subject.

3.2.3 Study 2: Whose speech children report

Whose speech do children report? Ely and McCabe (1993) found that English-speaking children present their own utterances in 53% of their spontaneous speech reports and the speech of non-parental adults, other children and their parents in the remaining 47%.

Interestingly, both English-speaking children and Spanish-speaking adults prefer direct speech when reporting their own utterance and indirect speech when reporting the utterance of someone else (Ely & McCabe, 1993; Van der Houwen, 2012). Can this difference between self and other also explain German children's choice of a direct or indirect speech report?

Table 8: Number (and percentage) of who is the reported speaker in German children's speech reports

	1S	2S	3S	1P	2P	3P
Direct	11 (14%)	3 (4%)	57 (75%)	0 (0%)	0 (0%)	5 (7%)
Indirect	5 (13%)	5 (13%)	26 (65%)	2 (5%)	0 (0%)	2 (5%)
Total	16 (14%)	8 (7%)	83 (72%)	2 (2%)	0 (0%)	7 (6%)

As table 8 shows, in the majority of the reports (72%), German children present the speech of a singular third person. Self-reports are the second most frequent report type (14%), while reports of other types of speakers are relatively rare.¹¹ If we compare these results to the results of previous studies, it is striking that German children produce considerably fewer self-reports than English-speaking children (14% vs. 53%) (Ely & McCabe, 1993). What is more, in contrast to findings from Ely and McCabe (1993) and Van der Houwen (2012), in our German corpus children do not show a strong preference for reporting their own speech with a direct speech report. In fact, self-reports are as frequent in direct speech as in indirect speech (14% vs. 13%).

These divergent results could be caused by differences in the data sets. Van der Houwen (2012) analyzed sociolinguistic interviews, in which the speakers were asked to talk about themselves. Ely and McCabe (1993) explicitly confined their analysis to reports of real persons' speech, excluding all cases in which children report the speech of their pets, dolls or imaginary friends. Our speech report corpus is not limited to reality-oriented reports, but includes also fictive interactions, which are an integral part of natural interactions (cf. Pascual, 2014).

¹¹ Similarly, Dutch children also show a clear preference for reporting the speech of a singular third person (n = 20, 59%).

To make the data basis more comparable to that of the previous studies mentioned above, I restrict the analysis in a next step to reports in which the reported speaker is a real person. As table 9 indicates, this leads to a drastic change.

Table 9: Number (and percentage) of who is the reported speaker in German children's speech reports, restricted to real persons

	1S	2S	3S	1P	2P	3P
Direct	10 (50%)	1 (5%)	9 (45%)	0	0	0
Indirect	4 (15%)	3 (12%)	17 (65%)	2 (8%)	0	0
Total	14 (30%)	4 (9%)	26 (57%)	2 (4%)	0	0

Now the majority of direct speech reports are self-reports (50%), while the number of indirect speech reports with the self as reported speaker is considerably lower (15%). When reporting the speech of other people, especially that of an individual third person, German children tend to use indirect speech. This is in line with previous findings that the self–other distinction influences the choice of report type (Ely & McCabe, 1993; Van der Houwen, 2012). Van der Houwen suggests that speakers prefer to use direct speech in self-reports because “it is fairly likely that the reporter will feel involved when talking about him/herself” (Van der Houwen, 2012, p. 123). However, apart from the fact that this is a rather obscure statement, the whole claim is weakened by the fact that the difference between self-reports and other-reports only seems to surface once we restricted our analysis to real persons. The reason for this is that children report the speech of fictive persons also mainly (80%) with direct speech reports.

3.2.4 Study 3: The influence of the discourse context

Introduction

A main finding of section 3.1 was that German and Dutch children produce much more direct than indirect speech reports in natural interactions. I hypothesized that this could be due to the discourse contexts in which the conversations are situated. Children and their interlocutors often engage in activities such as story-telling and role-playing, which are focused on fun and entertainment – perlocutionary effects typically associated with direct speech. The prediction is that in other types of discourse contexts, in which the transmission of propositional content is central, children might use more indirect speech (Clark & Gerrig, 1990; Tannen, 2007; Wade & Clark, 1993).

Previous studies are largely in line with these considerations. Nordqvist (2001a) analysed in detail how different discourse contexts (make-belief play, book reading, personal narratives, and habitual & hypothetical topics) influence the reports of Swedish children's parents. Her findings show that adults adapt their reports to different discourse contexts. While Swedish adults preferred direct speech in all discourse contexts, the proportion of indirect speech was highest when they talked about habitual or hypothetical topics (Nordqvist, 2001a). Vincent and Perrin (1999) found that French-speaking adults used more direct or indirect speech depending on the function of the report: Indirect speech is associated with an authority function, whereas direct speech is associated with a narrative function.

In contrast to adults, it is still an open question whether children's use of direct and indirect speech is affected by the discourse context or the communicative function of the report. Previous production studies indicate similar tendencies as in adults. Children produce, for instance, predominantly direct speech reports in narratives and play contexts (Ely & McCabe, 1993; Nordqvist, 2000; Özyürek, 1996) and indirect speech to tell on others or negotiate behaviour with authorities (Boeg Thomsen, 2014).

However, these findings have to be interpreted with caution because they originate from different groups of children, different methodologies (natural recordings vs. elicitation) and different languages. In order to show that children actually adjust their reports to the surrounding discourse context and the function of the report, data of the same group of children across different conversational contexts needs to be analyzed. In the following, I will provide such an analysis.

Method

I classified all German speech reports in the pronoun corpus in two respects: (i) the discourse context in which they were produced, and (ii) whether or not they express an authority function. I will explain the classifications in detail below.

Discourse contexts. I distinguish three types of discourse contexts: narrative, pretend play and reality. In the following, I point out the characteristics of each of these discourse contexts and illustrate them with an example.

The discourse context *narrative* subsumes situations of story-telling and book-reading. As an example of a narrative context, consider the following situation. A three-year-old girl and a friend of the family look together at pictures of the book "Der Struwwelpeter" (1845) by Heinrich Hoffmann. The adult prompts the child to tell the story of Suppenkaspar, a boy who refuses to eat and subsequently withers away and dies. Complying with the request, the child reports a characteristic utterance of Suppenkaspar:

- (24) da sagt der: „ich esse meine ess nich(t) die Suppe“ (3;1)
‘then he said: “I eat my don’t eat the soup”’

The second type of discourse context, *pretend play*, covers different kinds of play activities, such as embodied role play or pretend play with stuffed animals, toy figures or dolls (cf. Harris, 2000). To give an example of a pretend play context, imagine the following situation. A three-year-old girl is playing with toy figures, cars and a multi-story play car park. She lets a car with a boy figure slide down the sloping driveway and reports the boy's utterance:

- (25) un dann sag' der: „hurra ich bin gerutscht“ (3;6)
 'and then he says: "hooray I was sliding"'

The category *reality* includes contexts that are situated in the 'real world' in contrast to the fictive worlds of narrative and pretend play. The situations in the 'reality' category are quite diverse, ranging from having lunch to playing board games. As an example, take the situation that a mother is preparing her children for going outside. In this context, her three-year-old son reports the utterance of an absent person:

- (26) die Martina hat gesagt, sie wollte in 'n Garten (3;5)
 'Martina said she wanted (to go) into the garden'

Authority function. Speech reports can have many different functions, as conversation analysts have shown (see 2.2.8). I focus here on one function in particular that seems to be especially important for children: the authority function. In their analysis of spontaneous speech reports, Ely and McCabe note:

Children's disposition to recall voices that describe who is to do what suggests that they have a keen concern with how authority and power are exercised in their daily lives. (Ely & McCabe, 1993, p. 688)

Following Vincent and Perrin (1999), I determine that a report has an authority function if the reporting speaker intends to communicate the content of an utterance and wants to increase its persuasive power by appealing to the reported speaker as authority. In example (27), for instance, a boy appeals to his father in an argument with his mother about whether or not he is allowed to jump from a certain ladder:

- (27) nur dass der Papa gesagt hat, dass ich das darf (3;6)
 'just that Daddy said that I am allowed to do that'

All speech reports in the German pronoun corpus were classified into two categories: reports that do and reports that do not express an authority function.

Results

Table 10 shows how often German children produce direct and indirect speech reports in the three types of discourse contexts that I distinguish. We find that children use direct speech mainly in narrative and pretend play contexts. Indirect speech reports, however, show a clear preference for reality contexts. As table 11 indicates, indirect speech reports have more often an authority function than direct speech reports.

Table 10: Number (and percentage) of direct and indirect speech reports in different discourse contexts

	Direct speech	Indirect speech
Narrative	28 (37%)	4 (10%)
Pretend play	33 (43%)	10 (25%)
Reality	15 (20%)	26 (65%)

Table 11: Number (and percentage) of direct and indirect speech reports with authority function

	Direct speech	Indirect speech
Authority function	3 (4%)	14 (35%)
Other function	73 (96%)	26 (65%)

In order to check whether discourse context and authority function – and in addition also children's age – predict the use of indirect speech, I used generalized linear mixed effects models with the software *R* (version 3.2.0). With a procedure of model comparison, I added the fixed-effect factors DISCOURSE CONTEXT, AUTHORITY FUNCTION and AGE stepwise to the baseline model (including random intercepts for subjects). A factor was included if it led to a decrease of the Akaike Information Criterion (AIC) of more than 2 (Akaike, 1974). We found main effects of all three factors. The index of concordance of the model is 0.84, which indicates that the model may have real predictive power (Baayen, 2008).

Table 12 shows the fixed-effects coefficients of the model that explains German children's use of indirect speech. As evident from the positive estimate, children are more likely to produce indirect speech in reality contexts than in narrative contexts ($p = .002$). The production of indirect speech in contexts of pretend play and narratives did not differ significantly ($p = .261$). When a report has an authority function, the chance that a speaker uses an indirect speech report is significantly increased ($p = .003$). In addition, with increasing age, children tend to use more indirect speech reports ($p = .007$).

Table 12: Fixed-effects coefficients of the model explaining German children's use of indirect speech

	Estimate	SE	z value	p value
(Intercept)	-6.060	1.598	-3.792	<.001
Play vs. Narrative	0.769	0.684	1.124	.261
Reality vs. Narrative	2.088	0.662	3.156	.002
AUTHORITY FUNCTION	2.362	0.785	3.009	.003
AGE	1.190	0.444	2.679	.007

Note. Model includes random intercepts for subjects (N = 13, age range: 2;10–4;6).

Discussion

The results of pilot study 3 provide evidence that children under the age of five are sensitive to the communicative functions of direct and indirect speech. They tend to produce the more vivid direct speech type in playful and entertaining contexts and the more content-oriented indirect speech type in reality contexts. These findings also support my hypothesis that the generally higher frequency of direct speech reports in our corpus could be caused by the predominance of playful situations.

I focused on one specific function of speech reports, the so-called authority function, which has previously been linked to indirect speech reports (Boeg Thomsen, 2014; Vincent & Perrin, 1999). Similar to Danish children and Canadian adults, German children prefer indirect speech when appealing to authorities.

3.3 SUMMARY OF SPONTANEOUS SPEECH ANALYSIS

In a corpus study, I have analyzed Dutch and German children's use of speech reports in natural interactions. The results show that Dutch and German children start to produce speech reports with the reporting verb *say* from two years on. Children's first reports are direct quotations of single words or sounds, such as *says "moo"*. Several months later, children also quote clauses with a finite verb and produce indirect speech reports. This order suggests that the development of reports is related to syntactic complexity. The syntactically more complex clausal direct speech and (verb-final and verb-second) indirect speech constructions are acquired later than quotations of simple linguistic and non-linguistic elements. The quantitative analysis has revealed that in our corpus both children and their interlocutors clearly prefer to use direct speech reports. This could be due to the fact that many interactions between children and their caretakers take place during book-reading and pretend play activities, which favor the more vivid report type, direct speech.

In three pilot studies, I have focused on direct and indirect speech reports with a pronoun in subject position of the reported utterance (e.g., *Anna said that she is happy*). The first pilot study showed that children have different pronominal preferences in direct and indirect

speech. This can be explained by the different deictic point of orientation in direct and indirect speech. The second pilot study revealed that children report mostly the speech of singular third persons. In the final pilot study, I found evidence that children's choice of a direct or indirect speech report depends on the discourse context and the function of the report.

The finding that pragmatic factors such as the discourse context influence children's production of speech reports, raises the question whether children's comprehension of direct and indirect speech depends on pragmatic factors as well. In the following three experiments (chapters 4, 5 and 6), I assess children's interpretation of pronouns in speech reports in different communicative situations. In experiments 1 and 2, the speech reports have an authority function. In experiment 3, participants listen to reports that are embedded in a fictional narrative.

4 EXPERIMENT 1: WHO GETS THE BALL? A TABLET GAME TO TEST REPORTED SPEECH COMPREHENSION¹

In the first experiment, I investigate the processing of deictic pronouns in direct speech (*Elephant said, “I get the football”*) and indirect speech (*Elephant said that I get the football*). In an interactive tablet game, participants watch short animations and have to select the referents of first-, second-, and third-person pronouns embedded in direct or indirect speech.

First, I examine whether it is cognitively more demanding for Dutch adults to interpret pronouns in direct or indirect speech. As it turns out, adults have higher error rates and longer decision times when interpreting pronouns in direct as opposed to indirect speech (4.1). Subsequently, I follow the development of Dutch children between 4 and 12 years. The results indicate that children interpret pronouns in direct speech predominantly like in indirect speech, supporting our hypothesis about a late acquisition of the direct–indirect distinction (4.2).

4.1 ADULTS' INTERPRETATION OF PRONOUNS IN SPEECH REPORTS

4.1.1 Introduction and hypotheses

An essential feature of language is that it allows us to talk about what others have said. For this specific purpose, many languages such as Dutch and English, offer two distinct constructions: direct speech (1a, 1b) and indirect speech (2a, 2b).

- (1) a. Anna zei, “Ik ben blij”. (Dutch)
- b. Anna said, “I am happy”. (English)
- (2) a. Anna zei dat ze blij is. (Dutch)
- b. Anna said (that) she is happy. (English)

Semantically, direct and indirect speech can be distinguished in terms of perspective (Clark & Gerrig, 1990; Coulmas, 1986). In indirect speech, the reported utterance is paraphrased from the reporting speaker's perspective, i.e., in *his* words. In particular, the reporter in (2) uses

¹ Chapter 4 is based on Köder, Maier, and Hendriks (2015) and Köder and Maier (2015).

a third-person pronoun to refer to Anna from his own perspective. By contrast, in direct speech the reporting speaker purports to present a certain utterance verbatim and therefore shifts to the perspective of the original speaker. Consequently, the first-person pronoun in (1) refers not to the actual speaker himself but to Anna. In sum, interpreting direct speech requires a shift from the reporting to the reported speaker's perspective, which surfaces linguistically as a shift in the interpretation of deictic elements, like the pronoun *I*. This perspective shift might make pronoun processing in direct speech more demanding than in indirect speech.

However, there is also evidence suggesting the contrary, i.e., that direct speech is the cognitively less demanding report type. While most psycholinguistic and neurolinguistic studies focus on the vividness associated with direct speech (Yao et al., 2011, 2012; Yao & Scheepers, 2011), Groenewold, Bastiaanse, Nickels, Wieling, and Huiskes (2014) investigated how the use of direct and indirect speech affects discourse comprehension. They found that Dutch speakers with and without aphasia understood stories better when they contained direct speech. They attribute the positive effect of direct over indirect speech to the greater liveliness and the syntactic simplicity of direct speech (viz., coordination rather than subordination, De Vries, 2008). Our finding in chapter 3 that Dutch and German children produce more direct than indirect speech in natural interactions provides additional support for the thesis that direct speech is less demanding.

In experiment 1, we focus on the interpretation of deictic pronouns in direct and indirect speech. Based on the above considerations, we put forward the following two hypotheses. The first hypothesis is that pronoun processing is harder in direct speech because it requires a cognitively demanding shift from the reporting to the reported speaker's perspective. The second hypothesis points in the opposite direction: pronouns in direct speech are easier to interpret than in indirect speech. This second hypothesis is motivated by the positive effect of direct speech on story comprehension and children's frequent use of direct speech reports. To test these hypotheses, we created a referent selection task, in which we measure the relative processing costs of direct and indirect speech in terms of error rate and decision time for the interpretation of pronouns.

A second question that we investigate is whether the type of pronoun (first-, second-, or third-person) influences the interpretation process. The deictic singular pronouns refer to three different participant-roles in a communicative interaction: first-person *I* refers to the speaker, second-person *you* to the addressee, and third-person *he*, *she* and *it* to other people and things. Among deictic pronouns, third-person pronouns are claimed to be fundamentally different from first- and second-person pronouns (Lyons, 1977; Schlenker, 2003). While *I* and *you* refer directly to the primary participants of a speech situation, the meaning of deictic third-person pronouns is negatively defined as referring to neither the speaker nor the addressee. In Kaplan's (1989) framework, first- and second-person pronouns are pure indexicals, directly getting their reference from the context parameters, while deictic third-person pronouns are genuine demonstratives, requiring a pointing to be interpretable. This semantic distinction is supported by developmental studies showing that children's comprehension of

third-person pronouns lags behind their comprehension of first- and second-person pronouns (Brenner, 1983; Charney, 1980; Legendre & Smolensky, 2012). We examine whether this split between first- and second- versus third-person pronouns also holds in indirect and especially direct speech reports, where their interpretation requires a perspective shift.

In sum, the present study investigates adults' interpretation of pronouns in speech reports. In particular, we examine the difference between two report types (direct speech, indirect speech) and three pronoun types (first-, second-, third-person).

4.1.2 Method

4.1.2.1 Participants

Participants of the study were 116 native speakers of Dutch (74 women, $M_{age} = 23.0$ years, age range: 18–61 years). Of these participants, 85 took part in an online version of the experiment running in a web browser, and 31 were tested by an experimenter on a touchscreen tablet. The test items for both groups were identical. The only difference was whether the instructions were prerecorded or presented by an experimenter. All participants were recruited by the digital learning environment and mailing lists from the University of Groningen.

4.1.2.2 Stimuli and Procedure

The experiment is designed as a game that participants played on a tablet or computer. It uses an Android application linked to a small offline webpage. JavaScript controls the scenes and interactive responses and times them using the system time in milliseconds. The experimental game is called *Who gets the ball?* and is about identifying which of three animals receives a certain object.²

The game starts with an introduction phase in which the three main protagonists, a dog, a monkey and an elephant, introduce themselves and it is checked whether the participants remember their names (Hond 'Dog', Aap 'Monkey', Olifant 'Elephant'). Each of the animals is voiced by a different male speaker of Dutch. In addition, 18 familiar objects (e.g., football, book, car) that are part of the game are presented and named.

The test phase is split in two parts: the 'no report' condition and the 'speech report' condition (see table 1). First 15 'no report' items were presented, followed by 30 'speech report' items (15 direct, 15 indirect), randomized within these two blocks. Every participant saw the test items in a different random order. While test sentences in the 'no report' condition have the form of simple non-embedded statements like *I get the car*, test sentences in the 'speech report' condition are either direct or indirect speech reports preceded by a reporting

² All test sentences can be found in Appendix A: Experiment 1. An online version of the game can be played at <http://tinyurl.com/o7bburc> (Google Chrome required).

clause such as *Monkey said*. Our Dutch direct and indirect speech stimuli are clearly distinct in several syntactic, lexical and prosodic respects. Direct speech sentences have verb-second word order in the report, indirect speech sentences have verb-final word order and include the complementizer *dat* ‘that’. In addition, direct speech sentences have an 800 ms break between reporting clause and quotation and include a change of pitch in the quotation.

Table 1: Overview of test conditions

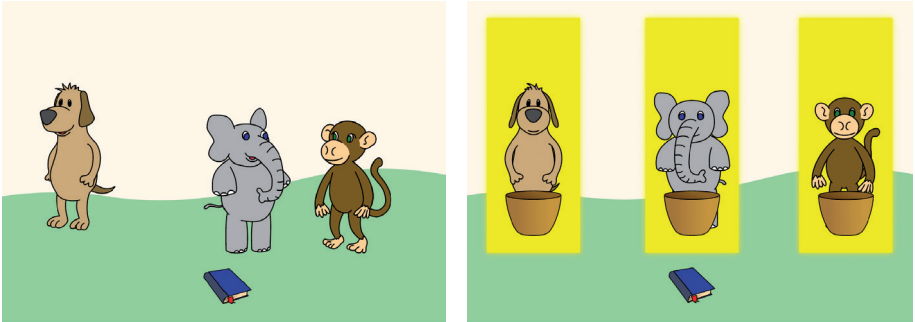
Condition	Test sentences
No report	Ik/ Jij/ Hij krijg(t) de/ het [object]. 'I/ You/ He get(s) the [object].'
Speech report	
a. Direct	Aap/ Olifant/ Hond zei, "Ik/ Jij/ Hij krijg(t) de/ het [object]." 'Monkey/ Elephant/ Dog said, "I/ You/ He get(s) the [object]."
b. Indirect	Aap/ Olifant/ Hond zei dat ik/ jij/ hij de/ het [object] krijg(t). 'Monkey/ Elephant/ Dog said that I/ you/ he get(s) the [object].'

All test sentences contain either a first-, second- or third-person singular pronoun (*ik* ‘I’, *jij* ‘you’ or *hij* ‘he’). Note that the third-person pronoun *he* is used deictically in the ‘no report’ condition and in direct speech, but anaphorically in indirect speech. In its deictic use, third-person *he* is usually accompanied by additional cues such as pointing or eye gaze (Kaplan, 1989; Stukenbrock, 2015). We decided against the inclusion of additional information in order to keep the third-person stimuli uniform with the first- and second-person ones. This could be at the expense of the naturalness of the deictic third-person pronouns in our experiment. Third-person *he* in indirect speech (*Monkey said that he gets the plane*) is special in our experiment. This is the only instance in which a pronoun is used anaphorically, i.e., to refer back to the subject of the matrix clause. As a consequence, the comparisons with indirect *he* are potentially confounded by the deictic–anaphoric distinction.

The test sentences are uttered in the context of communicative interactions between the three animals. We opted for the use of animations instead of static pictures. This allows us (i) to simulate natural interactions more closely, for instance by mimicking the mouth movement of the speaker, and (ii) to create a more engaging environment for children. To give an impression what the game looks like, we describe in more detail an example of a ‘no report’ item and of a ‘speech report’ item.

In the ‘no report’ condition, each animated scene involves the following actions. The elephant, for instance, walks over to the monkey and tells him who gets the book by uttering the sentence *He gets the book* (fig. 1a). After the utterance, all three animals are highlighted in yellow and a basket appears in front of them (fig. 1b). The participants’ task is to select the recipient of the book by touching or clicking on him. In our example, the correct choice would be the dog. After selection, the object jumps into the basket of

the chosen animal. The software records the accuracy of pronoun interpretation and the decision time in milliseconds from the offset of the sentence. Note that decision time in our experiment is not a measure of online processing, but an indication of the time participants need to make a decision after the whole sentence has been presented.



(a) Uttering of a 'no report' sentence (e.g., *He gets the book.*) (b) Selection phase

Figure 1: Example of a 'no report' item

To make sure that participants understand the procedure, we presented three practice items prior to the test items. They included the names of the animals instead of pronouns, for example, *Elephant gets the book.*

In the 'speech report' condition, the described interaction between the animals includes an additional step. Now one animal, for instance the elephant, walks over to the monkey and whispers into his ear who gets the object (fig. 2a). Participants heard only an incomprehensible whispering sound. Subsequently, the monkey walks to the dog and tells him what the elephant has said using either a direct or indirect speech report (fig. 2b). If the monkey says, for instance, *Elephant said, "I get the football"*, the correct referent of the pronoun *I* in this direct speech report is the speaker of the reported utterance, that is, the elephant. In contrast, the referent of *I* in an indirect speech report such as *Elephant said that I get the football* is the reporting speaker, the monkey.

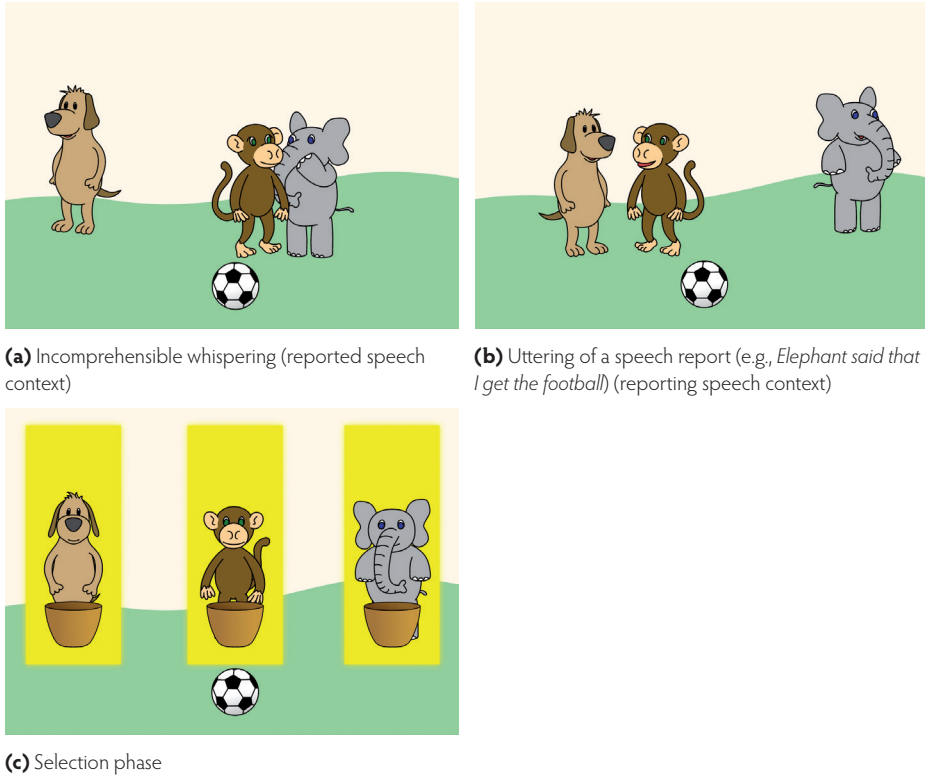


Figure 2. Example of a 'speech report' item

More generally, pronouns in direct speech need to be evaluated with respect to the reported context, that is the context in which the reported utterance was originally produced (Elephant whispering into Monkey's ear) (see fig. 2a). In indirect speech, pronouns need to be evaluated with respect to the reporting context, that is the context in which the monkey reports the elephant's utterance to the dog (see fig. 2b). Participants could identify the speaker (referent of *ik*) by his mouth movement and characteristic voice. The addressee (referent of *jij*) is turned towards the speaker, and the 'other person' (referent of *hij*) is positioned at a distance from speaker and addressee facing another direction. All animals have the same male gender, so that the gender feature on the third-person pronoun does not serve as an additional cue.

Note that direct and indirect speech reports appeared in random order within the 'speech report' condition and not in blocks. This means that for every speech report participants had to detect the direct or indirect speech cues in order to determine the appropriate interpretation of the pronoun. Experimental participants were external observers of the interaction and not possible referents of pronouns themselves. We counterbalanced the

communicative roles that the animal protagonists assume in the scenes (speaker, addressee, other), their spatial position and the types of sentences they utter. The experiment took approximately 10 to 15 minutes to complete.

4.1.3 Results

4.1.3.1 Accuracy

Figure 3 shows the percentage of Dutch adults' correct pronoun interpretation in non-reportative sentences and in direct and indirect speech reports. In the 'no report' condition, participants interpreted 100% of first- and second-person pronouns correctly, but only 63% of third-person pronouns. To focus on the direct–indirect distinction, our statistical analysis is restricted to the 'speech report' condition.³

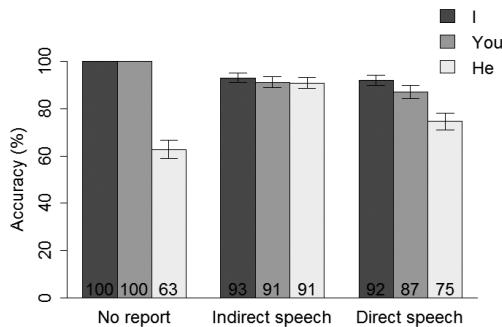


Figure 3: Percentage of correct pronoun interpretation in Dutch adults. Error bars indicate 95% confidence intervals.

We used mixed-effects logistic regression modeling with the software *R* (version 3.1.1) to analyze the accuracy of pronoun interpretation in direct and indirect speech. The random-effects structure of our model includes random intercepts for subjects and items and by-subject random slopes per REPORT TYPE. Step by step, we tested whether the following fixed-effect factors improve the goodness of fit of the accuracy model: REPORT TYPE (direct speech, indirect speech), PRONOUN TYPE (*I*, *you*, *he*), TEST CONDITION (online, tablet), SPEAKER (Dog, Monkey, Elephant), AGE and GENDER of the participants, EXPERIENCE (1–5) and SEQUENCE NUMBER (1–30). The factors EXPERIENCE and SEQUENCE NUMBER both in-

³ Note that a direct comparison between the results of the 'no report' and the 'speech report' condition would be complicated by differences in the scenarios and the fact that the 'no report' condition always preceded the 'speech report' condition.

dicating the progress in the experiment. `SEQUENCE NUMBER` provides information about the order in which a participant saw the ‘speech report’ items, `EXPERIENCE` indicates how often a participant has encountered a certain type of test item (e.g., direct *I*) before.

Table 2 shows that the goodness of fit of the accuracy model improves significantly by adding `REPORT TYPE` and `PRONOUN TYPE` as fixed-effect factors as well as their interaction. `EXPERIENCE` is also a significant predictor and has a positive effect on accuracy, $\beta = 0.49$, $z = 7.70$, $p < .001$. All other factors did not contribute to the model fit.

Table 2: Goodness of fit of the fixed-effect factors of the accuracy model for speech reports. An Akaike Information Criterion (AIC) decrease of more than 2 indicates that the goodness of fit of the model improves significantly (Akaike, 1974).

	AIC decrease
Random-effect factors (random intercepts for subjects and items, by-subject random slopes per <code>REPORT TYPE</code>)	
+ <code>REPORT TYPE</code>	12.2
+ <code>PRONOUN TYPE</code>	28.5
+ <code>REPORT TYPE * PRONOUN TYPE</code>	13.6
+ <code>EXPERIENCE</code>	64.5

We compared the mean accuracy of different combinations of report type and pronoun type using multiple comparisons from the ‘multcomp’ package (Hothorn, Bretz, & Westfall, 2008). The results are presented in table 3. The general tendency is that participants made significantly more mistakes when the same pronoun occurred in direct as compared to indirect speech (for *you*: $p < .05$, for *he*: $p < .001$). The first-person pronoun was an exception with similar accuracy rates in direct and indirect speech. In indirect speech, there were no significant differences between pronouns. In direct speech, the accuracy of *I* was significantly higher than that of *you* ($p < .01$) and *he* ($p < .001$). In turn, participants were more accurate in interpreting direct *you* than direct *he* ($p < .001$).

Table 3: Multiple comparisons of means for accuracy of pronoun interpretation in speech reports (Tukey contrasts)

Linear hypotheses	Estimate	SE	z value	p value
Indirect <i>I</i> – Direct <i>I</i> = 0	3.58	1.37	2.62	0.064
Indirect <i>You</i> – Direct <i>You</i> = 0	3.87	1.32	2.92	0.027
Indirect <i>He</i> – Direct <i>He</i> = 0	5.49	1.30	4.24	<0.001
Indirect <i>I</i> – Indirect <i>You</i> = 0	0.81	0.39	2.07	0.237
Indirect <i>I</i> – Indirect <i>He</i> = 0	0.93	0.39	2.40	0.112
Indirect <i>You</i> – Indirect <i>He</i> = 0	0.13	0.36	0.36	0.999
Direct <i>I</i> – Direct <i>You</i> = 0	1.09	0.29	3.82	0.001
Direct <i>I</i> – Direct <i>He</i> = 0	2.85	0.29	9.66	<0.001
Direct <i>You</i> – Direct <i>He</i> = 0	1.75	0.24	7.22	<0.001

A closer examination of the errors reveals an interesting pattern: in 94% of the mistakes in direct speech, participants interpreted the pronouns as in indirect speech. For instance, when Monkey whispers to Dog, and Dog then reports to Elephant *Monkey said*, “*You get the rose*”, some participants incorrectly chose Elephant as the referent of quoted *you* instead of Dog. This means that they evaluated pronouns in direct speech with respect to the reporting instead of the reported speech context, i.e., they did not perform the perspective shift necessary in direct speech. No such clear error pattern was found in indirect speech. In only 64% of the indirect speech mistakes (50% chance level) did participants wrongly select the direct speech referent.

4.1.3.2 Decision times

The decision times for different conditions and pronouns are presented in figure 4. Our analysis focuses on comparing participants’ decision times for direct and indirect speech reports. We fitted a linear mixed-effects model to the log-transformed decision times. Using a procedure of model comparison, we added stepwise fixed-effect factors to the baseline model (with random intercepts for subjects and items, and by-subject random slopes per REPORT TYPE).

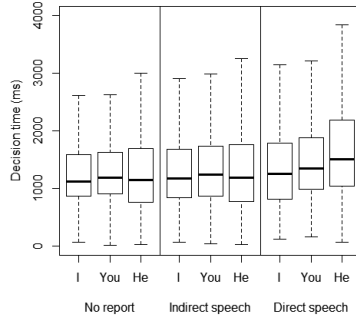


Figure 4. Boxplots of decision times (excluding outliers).

Table 4 shows that the inclusion of REPORT TYPE and PRONOUN TYPE and their interaction improves the goodness of fit of the model significantly. Other significant predictors are EXPERIENCE and TEST CONDITION (online, tablet). The more experience participants had with a certain type of test item, the quicker they were in their response, $\beta = -0.08$; $t = -13.27$; $p < .001$. Participants with the online browser version had shorter decision times than those with the tablet version, $\beta = -0.40$; $t = -5.79$; $p < .001$.

Table 4: Goodness of fit of the fixed-effect factors of the model fitted to log-transformed decision times of speech reports. An Akaike Information Criterion (AIC) decrease of more than 2 indicates that the goodness of fit of the model improves significantly (Akaike, 1974).

	AIC decrease
Random-effect factors (random intercepts for subjects and items, by-subject random slopes per REPORT TYPE)	
+ REPORT TYPE	14.4
+ PRONOUN TYPE	3.3
+ REPORT TYPE * PRONOUN TYPE	8.8
+ EXPERIENCE	169.4
+ TEST CONDITION	26.4

Table 5: Multiple comparisons of means for decision times of speech reports (Tukey contrasts)

Linear hypotheses	Estimate	SE	z value	p value
Indirect <i>I</i> – Direct <i>I</i> = 0	-0.01	0.05	-0.32	1.000
Indirect <i>You</i> – Direct <i>You</i> = 0	-0.17	0.05	-3.55	0.005
Indirect <i>He</i> – Direct <i>He</i> = 0	-0.27	0.05	-5.80	<0.001
Indirect <i>I</i> – Indirect <i>You</i> = 0	-0.00	0.04	-0.06	1.000
Indirect <i>I</i> – Indirect <i>He</i> = 0	0.02	0.04	0.37	0.999
Indirect <i>You</i> – Indirect <i>He</i> = 0	0.02	0.04	0.43	0.998
Direct <i>I</i> – Direct <i>You</i> = 0	-0.16	0.04	-3.53	0.006
Direct <i>I</i> – Direct <i>He</i> = 0	-0.24	0.04	-5.51	<0.001
Direct <i>You</i> – Direct <i>He</i> = 0	-0.09	0.04	-1.98	0.351

Table 5 shows the results of a multiple comparison analysis that compares the decision times for different items. It took participants significantly longer to interpret the pronouns *you* and *he* in direct speech as compared to indirect speech (for *you*: $p < .01$, for *he*: $p < .001$). The decision times for *I* did not differ with report type. While all pronouns were interpreted equally fast in indirect speech, participants needed more time to select the referents of *you* ($p < .01$) and *he* ($p < .001$) as compared to *I* in direct speech.

4.1.4 Discussion

In a referent selection task, we examined the interpretation of pronouns in direct and indirect speech. We found that Dutch adults made significantly more mistakes and had longer decision times when interpreting the pronouns *jij* ‘you’ and *hij* ‘he’ in direct speech as compared to indirect speech.

This finding supports our first hypothesis, that the perspective shift in direct speech increases the processing effort of pronouns. To further explain why perspective shifting would be so cognitively demanding, we follow a suggestion from the theory of mind literature, where perspective shifting difficulties in false belief tasks are explained in terms of the cost of inhibiting one’s own current perspective (Carlson & Moses, 2001; Saxe, 2009; Wellman, Cross, & Watson, 2001). Thus, we suggest that a proper interpretation of direct speech requires the listener to inhibit his representation of the actual reporting speaker’s perspective. Since, at the time of the report, this primary perspective is more salient than the reported speaker’s secondary perspective, inhibiting it is costly and error-prone. This explanation is supported by the fact that in 94% of the errors in direct speech, pronouns were evaluated with respect to the reporting speaker as deictic center, meaning that this primary perspective was not properly inhibited.

Participants performed surprisingly well on the first-person pronoun *I* in direct speech, which approximates the accuracy rates and decision times of indirect speech. A possible explanation for the apparent ease of direct *I* is that its referent is explicitly mentioned

in the reporting clause and hence a much more salient candidate. This explanation may be tested by checking whether performance on the second-person pronoun improves when we explicitly add the original addressee to the reporting clause, as in *Elephant said to me, "You get the football"*.

The results of our comprehension experiment seem to conflict with previous findings that indicate that direct speech is the cognitively easier report type. The corpus study in chapter 3 showed that Dutch children produce more direct than indirect speech in natural interactions. Moreover, there is evidence that Dutch adults with and without aphasia understand narratives containing direct speech better than narratives containing indirect speech (Groenewold et al., 2014). How can this be reconciled with the results of the current experiment that pronouns in direct speech are more difficult to process than in indirect speech?

This discrepancy can be dissolved if we differentiate between different discourse contexts in which the reports are embedded. In the current experimental game, reports were used to convey the information who gets a certain object. The reporting speaker uses a reporting clause (e.g., *Dog said*) in order to appeal to the person from whom he has received this information. The reports in this experiment therefore have an authority function, which has previously been shown to favor the use of indirect speech (Vincent & Perrin, 1999; this dissertation, section 2.2.4). By contrast, in Groenewold et al.'s (2014) study, the direct and indirect speech reports were presented as part of a narrative. Narrative contexts are associated with a predominant use of direct speech, allowing for more vivid storytelling (Vincent & Perrin, 1999; this dissertation, section 2.2.4).

This suggests that the type of discourse context does not only influence speakers' production of speech reports, but also listeners' comprehension. The increased processing effort associated with the perspective shift in direct speech explains why, in the information-exchange context of our experiment, the use of direct speech is clearly dispreferred, and in the case of direct *he* even pragmatically awkward. However, in a narrative context, the costs of a perspective shift in direct speech might be reduced because in a narrative the reported speaker's (i.e., the character's) perspective is generally more salient than the reporting speaker's (i.e., the narrator's) perspective.

Our second question concerned the difference between the interpretation of first-, second- and third-person pronouns. In the 'no report' condition as well as in direct speech, Dutch adults made most mistakes in the interpretation of *he*. This provides empirical support for the claim that there is a split between first- and second- versus third-person deictic pronouns (Lyons, 1977). The high number of mistakes with *he* in the 'no report' condition could be due to the lack of an accompanying pointing gesture or eye gaze. As pointed out in the method section, deictic *he* typically requires a pointing gesture to single out a unique referent (Diessel, 2012; Kaplan, 1989; Levinson, 2004). We decided against including pointing as it would trivialize the interpretation of *he* and break the uniformity of the test items across conditions and pronoun types. Note that this complication does not extend to the crucial 'speech report' condition because pointing gestures are not expected to be reported.

This explains the higher accuracy rate of *he* in direct and indirect speech reports as compared to non-embedded statements. In indirect speech, accuracy and decision times did not differ between *he* and the other pronouns. However, in indirect speech *he* is used anaphorically rather than deictically.

An anonymous reviewer of the Journal *Language, Cognition and Neuroscience* suggested that the deictic–anaphoric distinction could explain our findings without appealing to perspective shifting. We concede that, for *he*, the deictic–anaphoric distinction is a confounding factor. However, the difference between deixis and anaphora cannot account for the significantly lower performance (both in accuracy and decision time) with direct *you* as compared to indirect *you*. Additional evidence in favor of our perspective shifting hypothesis is the systematic error pattern pointed out above: Almost all direct speech errors, across pronoun type, involve a failure to shift from the actual to the reported speaker’s perspective.

In conclusion, this study provides evidence that the perspective shift in direct speech increases the processing effort of deictic pronouns in speech reports. Further studies are necessary to examine whether this effect can be replicated for other deictic expressions (e.g., spatial and temporal indexicals like *here* and *tomorrow*) and other discourse contexts (e.g., narratives).

4.2 CHILDREN’S INTERPRETATION OF PRONOUNS IN SPEECH REPORTS

4.2.1 Introduction

The previous section (4.1) has shown that Dutch adults make more mistakes and have longer decision times when interpreting the pronouns *you* and *he* in direct speech as opposed to indirect speech. This suggests that the perspective shift in direct speech increases the processing effort of pronouns. How will Dutch children perform in the same experiment? Children do not only have attested difficulties in shifting perspective (Gopnik & Astington, 1988; Masangkay et al., 1974; Wimmer & Perner, 1983), but more fundamentally they also do not seem to strictly distinguish between direct and indirect speech, according to previous studies (see section 2.4 for a detailed overview).

In the production of speech reports, children sometimes mix features of direct and indirect speech in an ungrammatical way. In (3), for instance, a child combines the complementizer *that*, a feature of indirect speech, with pronouns presented from the reported speaker’s perspective, a feature of direct speech.

- (3) And the birdie_i said that I_i feel much prettier knowing how long my_i beak is. (age 8;6) (Goodell & Sachs, 1992, p. 407)

However, while the elicitation studies of Goodell and Sachs (1992) and Hickmann (1993) found many cases of direct–indirect mixes, the corpus study in chapter 3 led to the surprising result that Dutch and German children rarely produce mixes of direct and indirect speech in natural interactions.

So far, little is known about children’s comprehension of speech reports. We know, for instance, that six-year-old children allow *wh*-extraction out of direct speech complements, a grammatical operation only admissible in indirect speech (Hollebrandse, 2007; Weissenborn et al., 1991). Furthermore, even twelve-year-olds show a non-adult-like ability to detect prosodic cues for direct and indirect speech (Hewlett et al., 2003). These two studies suggest that the correct interpretation of direct and indirect speech has a protracted development.

In experiment 1, we use the semantic difference in pronoun interpretation to find out when children can reliably tell apart a direct from an indirect speech report in Dutch. In order to interpret pronouns in speech reports correctly, it is not sufficient that children know the meaning of pronouns (e.g., that *ik* ‘I’ refers to the speaker), they also need to evaluate the pronouns with respect to the correct speech context.

- (4) a. Anna zei: “Ik ben blij”.
 ‘Anna said: “I am happy”’
 b. Anna zei dat ik blij ben.
 ‘Anna said that I am happy.’

In the direct speech report (4a), *ik* refers to the reported speaker Anna. But in the indirect speech report (4b), *ik* picks out the actual reporting speaker. Selecting the correct referent of pronouns in speech reports therefore requires being sensitive to features such as word order and prosody that indicate whether a given report is presented in direct or indirect speech.

4.2.2 Hypotheses

We put forward two hypotheses: the first and main hypothesis is related to the direct–indirect distinction; the second hypothesis concerns the difference between first-, second-, and third-person pronouns.

The main hypothesis is that Dutch children fail to clearly distinguish direct and indirect speech, despite clear syntactic, lexical and prosodic cues. Children’s lack of a rigid direct–indirect distinction is expected to result in a systematic misinterpretation of pronouns. Instead of consistently evaluating pronouns in direct speech with respect to the reported context and pronouns in indirect speech with respect to the reporting context, we expect them to mix up these two contexts of evaluation. More specifically, we predict that children will be especially prone to interpreting direct speech as indirect speech. This is because interpreting direct speech requires a cognitively demanding perspective shift from the reporting to the reported speech context (cf. adults’ results in section 4.1).

The second hypothesis concerns differences in the interpretation of first-, second- and third-person singular pronouns. We predict that third-person pronouns are more difficult to interpret than first- and second-person pronouns. This prediction is based on a number of previous studies that found that children understand first- and second-person pronouns before third-person pronouns (Brener, 1983; Charney, 1980; Werner Deutsch & Pechmann, 1978; Legendre & Smolensky, 2012; Murphy, 1986). This time lag could be caused by a semantic difference between so-called local pronouns (*I, you*) and third-person pronouns (*he, she*). While local pronouns refer directly to the primary participants of a speech act – speaker and addressee –, third-person pronouns trigger the negative presupposition that they denote neither speaker nor addressee (Legendre & Smolensky, 2012; Lyons, 1977; Schlenker, 2003). We expect to find a similar first- and second- versus third-person split when pronouns are embedded in a direct or indirect speech report, causing more errors and longer decision times for *he* than for *I* and *you*.

4.2.3 Method

4.2.3.1 Participants

The participants of this study were 136 monolingual Dutch-speaking children between 4;1 and 12;8 years of age (see table 6). Two additional children were tested, but had to be excluded due to inattention of the child (1) or experimenter error (1). The participating children were recruited from three elementary schools in the North of the Netherlands. Written parental consent was obtained prior to the experiment. Children received a small reward (a sticker for younger children, a pen for older children) for participating. In addition, 33 native speakers of Dutch – mostly students – participated without compensation. The data of these adults has been discussed as part of a larger population in the previous section. Here they serve as an adult control group. All participants were tested individually in a quiet room at the school or university and played the experimental game on a tablet.

Table 6: Participants

Age group	Mean age	Range	Number	Gender (f/m)
4	4;6	4;1–4;11	20	12/8
5	5;8	5;5–6;0	23	13/10
6	6;6	6;1–7;0	29	14/15
7	7;8	7;1–8;6	25	13/12
9	9;11	9;7–10;5	20	9/11
11	11;11	11;4–12;8	19	9/10
adults	22	18–32	33	21/12

4.2.3.2 Stimuli and Procedure

The same referent-selection task as in section 4.1 was used (see there for a detailed description). On a tablet, participants watched animated scenes in which the animal protagonists uttered sentences in direct speech (e.g., *Aap zei, “Jij krijgt de auto”* ‘Monkey said, “You get the car”’) or indirect speech (e.g., *Aap zei dat jij de auto krijgt* ‘Monkey said that you get the car’). The participants had to select the referents of first-, second-, and third-person singular pronouns by touching one of the animals on the screen. An experimenter was present throughout. She explained the experimental game and made sure that the children kept their focus of attention on the test items.⁴

4.2.4 Results

We first present the results of children’s pronoun interpretation in non-reportative sentences and subsequently in direct and indirect speech reports.

4.2.4.1 ‘No report’ condition: Accuracy

Figure 5 shows the percentage of correct reference assignment for the pronouns *ik* ‘I’, *jij* ‘you’ and *hij* ‘he’ in the ‘no report’ condition (e.g., *He gets the book*). While even the youngest children are around ceiling for the comprehension of first- and second-person pronouns, their comprehension of third-person pronouns clearly lags behind. Eleven-year-old children show a correct interpretation of *hij* in only 33% of the cases, adults in only 59%. A detailed analysis of the incorrectly interpreted third-person pronouns reveals the following error pattern: In 98% of the errors, participants selected the addressee instead of the ‘other person’ as referent of *hij*.

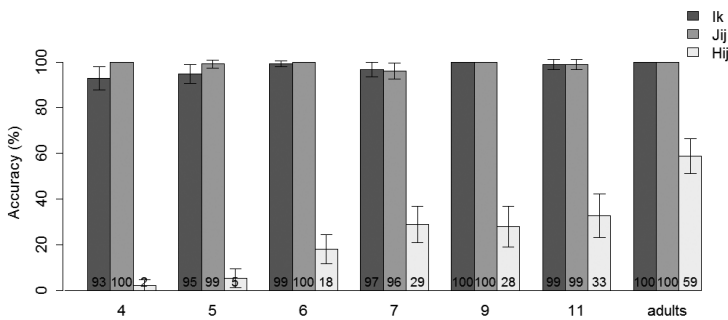


Figure 5: Percentage of correct pronoun interpretation in non-reportative sentences per age group. Error bars indicate 95% confidence intervals.

⁴ All test sentences can be found in Appendix A: Experiment 1. An online version of the game can be played at <http://tinyurl.com/o7bburc> (Google Chrome required).

We analyzed children's responses in the 'no report' condition with mixed-effects logistic regression modeling with the software *R* (version 3.1.1). Our baseline model includes random intercepts and slopes per PRONOUN TYPE for subjects, taking into account that participants might vary systematically in how they interpret the pronouns *ik*, *jij* and *hij*. Step by step, we tested whether the following factors improve the goodness of fit of the model: PRONOUN TYPE (*ik*, *jij*, *hij*), EXPERIENCE (1–5, indicating how many times a participant has seen an item with the same pronoun), SEQUENCE NUMBER (1–15, indicating how many 'no report' items a participant has already seen), SPEAKER (Monkey, Dog, Elephant), AGE (in months) and GENDER of the participants. The factors EXPERIENCE and SEQUENCE NUMBER are related. But whereas SEQUENCE NUMBER indicates the progress in the experiment, EXPERIENCE provides information about the content of the presented items, for instance, how often a specific participant has encountered the pronoun *ik* before.

A fixed-effect factor or an interaction was included in the model if it contributed significantly to the model fit as indicated by an Akaike Information Criterion (AIC) decrease of more than 2 (Akaike, 1974). Based on this procedure, our model for accuracy of pronoun interpretation in the 'no report' condition includes the fixed-effect factors PRONOUN TYPE and EXPERIENCE (see table 7). The index of concordance of the model is 1.00, which indicates that it has real predictive power (Baayen, 2008). Children made significantly more mistakes for the pronoun *hij* than for *ik* ($p < .001$), with no significant difference between *jij* and *ik* ($p = .963$). EXPERIENCE has a positive effect on accuracy, as evident from the positive estimate.

Table 7: Fixed-effects coefficients of the model fitted to children's accuracy of pronoun interpretation in 'no report' condition

	Estimate	SE	z value	p value
(Intercept)	9.3688	1.7674	5.301	<.001
<i>jij</i> 'you' vs. <i>ik</i> 'I'	0.1331	2.8483	0.047	.963
<i>hij</i> 'he' vs. <i>ik</i> 'I'	-22.1586	2.1274	-10.416	<.001
EXPERIENCE	0.6978	0.1736	4.019	<.001

Note. Model includes random intercepts and slopes per PRONOUN TYPE for subjects.

4.2.4.2 'No report' condition: Decision times

We excluded all decision times longer than 10 s because after that time a replay button appeared and children could see the test item again if they had not made a choice yet. This leads to the exclusion of 10 data points (0.005% of the data).

On average, children chose a referent for the first-person pronoun *ik* in 1.82 s ($SD = 0.95$), for the second-person pronoun *jij* in 1.78 s ($SD = 0.94$) and for the third-person pronoun *hij* in 1.89 s ($SD = 1.09$). The model that we fitted to children's log-transformed

decision times in the ‘no report’ condition includes random intercepts for subjects and as fixed-effect factors PRONOUN TYPE and SEQUENCE NUMBER. It takes children significantly longer to select a referent for a third-person pronoun than for a first-person pronoun in non-reportative sentences ($\beta = 0.19$; $t(1912) = 3.97$; $p < .001$). There is no significant difference in decision times between first- and second-person pronouns ($\beta = -0.00$; $t(1913) = -0.02$; $p = .987$). Participants reacted faster the more of the 15 ‘no report’ items they have already seen (SEQUENCE NUMBER) ($\beta = -0.02$; $t(1910) = -4.17$; $p < .001$). We also found an interaction of PRONOUN TYPE and SEQUENCE NUMBER. The model predicts that the acceleration effect is bigger for *hij* than for *ik* ($\beta = -0.02$; $t(1916) = -2.97$; $p = .003$).

4.2.4.3 ‘Speech report’ condition: Accuracy

Figures 6 and 7 show the percentage of correct pronoun interpretation in direct speech (fig. 6) and indirect speech (fig. 7).

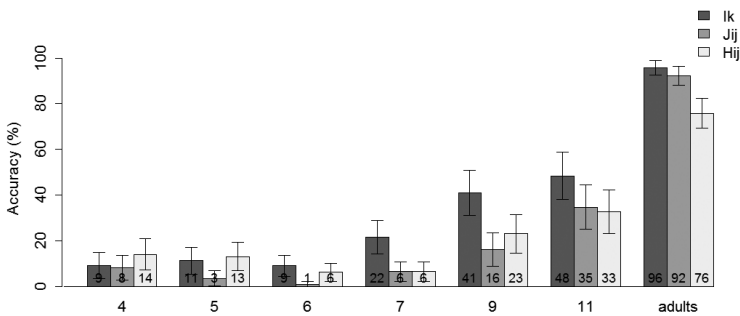


Figure 6: Percentage of correct pronoun interpretation in direct speech per age group. Error bars indicate 95% confidence intervals.

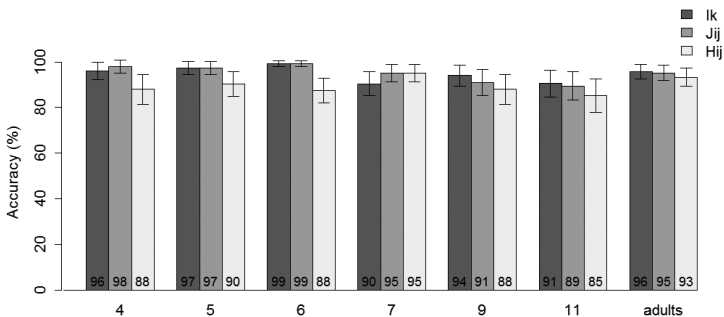


Figure 7: Percentage of correct pronoun interpretation in indirect speech per age group. Error bars indicate 95% confidence intervals.

We analyzed children's accuracy of pronoun interpretation in speech reports with mixed-effects logistic regression modeling, following a procedure of model comparison as described for the 'no report' condition. The best-fitted model contains as random-effect factors random intercepts and random slopes per REPORT TYPE for subjects and as fixed-effect factors REPORT TYPE (direct speech, indirect speech), PRONOUN TYPE (*ik*, *jjj*, *hij*) and AGE. In addition, we found a significant interaction between REPORT TYPE and PRONOUN TYPE. AGE has a positive effect on accuracy of pronoun interpretation ($\beta = 0.25$, $z = 4.76$, $p < .001$), meaning that older children made significantly fewer mistakes than younger children. The inclusion of all other factors (EXPERIENCE, SEQUENCE NUMBER, SPEAKER, GENDER) did not improve the goodness of fit of the model. The model has predictive power ($C = 0.98$).

Table 8: Multiple comparisons of means (Tukey contrasts) for children's accuracy of pronoun interpretation in the 'speech report' condition

Linear hypotheses	Estimate	SE	z value	p value
Indirect <i>ik</i> – Direct <i>ik</i> = 0	7.05	.55	12.81	<.001
Indirect <i>jjj</i> – Direct <i>jjj</i> = 0	8.68	.58	15.01	<.001
Indirect <i>hij</i> – Direct <i>hij</i> = 0	6.58	.52	12.76	<.001
Indirect <i>ik</i> – Indirect <i>jjj</i> = 0	-0.20	.32	-0.63	.986
Indirect <i>ik</i> – Indirect <i>hij</i> = 0	1.24	.27	4.56	<.001
Indirect <i>jjj</i> – Indirect <i>hij</i> = 0	1.44	.29	5.05	<.001
Direct <i>ik</i> – Direct <i>jjj</i> = 0	1.43	.20	6.96	<.001
Direct <i>ik</i> – Direct <i>hij</i> = 0	0.77	.18	4.26	<.001
Direct <i>jjj</i> – Direct <i>hij</i> = 0	-0.65	.21	-3.08	.019

We compared the means of different combinations of REPORT TYPE and PRONOUN TYPE across age groups with multiple comparisons (Tukey contrasts) from the 'multcomp' package, version 1.3–6 (Hothorn et al., 2008). The results are presented in table 8. Children made fewer mistakes when the same pronoun occurs in indirect speech in contrast to direct speech (for *ik*: $p < .001$, for *jjj*: $p < .001$, for *hij*: $p < .001$). In indirect speech, *hij* is the most difficult pronoun with a significantly lower accuracy than that of both *ik* ($p < .001$) and *jjj* ($p < .001$). The mean accuracy of *ik* and *jjj* in turn does not differ in indirect speech ($p = .986$). In direct speech, the first-person pronoun *ik* has the highest accuracy, significantly higher than that of *jjj* ($p < .001$) and *hij* ($p < .001$). Surprisingly, children were better in selecting the correct referent for third-person *hij* than for second-person *jjj* in direct speech ($p = .019$). However, if we look at all age groups individually, we find this tendency only in the data of four-, five-, six-, and nine-year-old children. In eleven-year-old children, this effect is reversed, meaning that the accuracy of second-person pronouns is higher than of third-person pronouns in direct speech. This is similar to what we find in adults.

4.2.4.4 'Speech report' condition: Decision times

For the decision times analysis, we excluded 17 data points that exceeded 10 s (0.004%) because after that time test items were presented anew. On average, children selected a referent in 1.90 s ($SD=1.04$) for direct speech reports and in 1.76 s ($SD=0.95$) for indirect speech reports. We fitted a linear mixed-effects regression model (with random intercepts and random slopes per REPORT TYPE for subjects and random intercepts for items) to children's log-transformed decision times. The model includes REPORT TYPE as fixed-effect factor. Participants have significantly shorter decision times when they hear an indirect as compared to a direct speech report ($\beta = -0.10$; $t(37.3) = -4.1$; $p < .001$). PRONOUN TYPE did not influence decision times in the 'speech report' condition.

4.2.4.5 Direct–indirect mixes

We call it a direct–indirect mix when children evaluate a pronoun in a report with respect to the incorrect speech context, i.e., in direct speech with respect to the reporting speech context and in indirect speech with respect to the reported speech context. Consider the example that the monkey utters the direct report *Elephant said*, “*I get the football*”. In this case a direct–indirect mix would be if the participant selected the monkey, the reporting speaker, instead of the elephant, the reported speaker, as the referent of *I*. 99% ($N=1721$) of the mistakes in direct speech are direct–indirect mixes. This deviates significantly from the chance level of 50% ($t(1720)=194.07$, $p < .001$). In indirect speech only 56% ($N=139$) of the mistakes are direct–indirect mixes, which is not significantly different from chance ($t(138)=1.45$, $p = .15$). This means that children predominantly interpret pronouns in direct speech like in indirect speech, but not the other way around.

4.2.5 Discussion

We formulated two specific hypotheses. The first concerned the acquisition of the direct–indirect distinction: Children will fail to distinguish direct and indirect speech and hence will make mistakes in interpreting pronouns embedded in speech reports. In particular, we predicted that they interpret pronouns in direct speech as if in indirect speech. The second hypothesis concerned children's interpretation of different types of pronouns: The third-person pronoun *hij* 'he' was expected to be harder to interpret than the local pronouns *ik* 'I' and *jij* 'you'. Both predictions were confirmed by the experiment. In this section we will discuss some *prima facie* puzzling patterns in the data.

Hypothesis 1: Children mix direct and indirect speech

The main hypothesis of our study is that children are not able to strictly distinguish direct and indirect speech in their interpretation of pronouns. Our data clearly confirm this, with a notable difference between direct and indirect reports. While even the youngest age group, the four-year-olds, was able to interpret pronouns in indirect speech correctly, children had much higher error rates and longer decision times for the interpretation of pronouns embedded in direct speech. This is in line with the results of Dutch adults that pronoun interpretation is cognitively more demanding in direct than in indirect speech. Children's accuracy of pronoun interpretation improved significantly with age. Yet, surprisingly, the performance of the oldest age group, the eleven-year-olds, was still not adult-like in direct speech, with a correct pronoun interpretation of only 48% for *I* (adults: 96%), 35% for *you* (adults: 92%) and 33% for *he* (adults: 76%).

A closer look at children's errors in direct speech reveals that 99% of the mistakes are direct–indirect mixes, that is, children evaluated pronouns in direct speech with respect to the reporting speech context, as in indirect speech. Mixes in the opposite direction are less frequent and turned out to be random. This strongly suggests that the reporting context is children's preferred context of evaluation for pronouns, regardless of report type.

Our findings are compatible with previous studies documenting children's difficulties with the direct–indirect distinction. From production studies we know that children up to the age of eight sometimes mix features of direct and indirect speech (Goodell & Sachs, 1992; Hickmann, 1993; Nordqvist, 2001a). Hollebrandse's (2007) comprehension study showed that five- and six-year-olds' interpretation of reports may simultaneously exhibit signs of *wh*-extraction, as in indirect speech, and pronoun shift, as in direct speech. Nonetheless, it is rather surprising to find that even eleven-year-old speakers of a language with such a clear marking of the direct–indirect distinction as Dutch still have not acquired that distinction. We propose three possible partial explanations for children's systematic misinterpretation of direct speech as indirect speech in our experiment.

The first explanation is that children are not sensitive to the prosodic features that, for adults, are an important direct speech cue. Indeed, Hewlett et al. (2003) showed that twelve-year-old children are less sensitive to prosody as a marker of the direct–indirect distinction than adults. However, this does not explain why children also ignored the unambiguous morphemic and syntactic direct speech cues (complementizer and word order) that we used in the construction of our Dutch stimuli.

The second explanation is that children have difficulties in dealing with perspective differences. Similar to false belief tasks (Wimmer & Perner, 1983) and alternative naming tasks (Perner, Stummer, Sprung, & Doherty, 2002), our task requires that children be able to represent the same entity from two different perspectives. In our experiment, one and the same animal assumes different participant-roles in the two speech contexts. The elephant can for instance be the addressee in the reported speech context and the speaker in the

reporting speech context. Children usually understand this type of perspective difference at around the age of four (Perner, Brandl, & Garnham, 2003). However, one could argue that the interpretation of pronouns in speech reports requires not first- but second-order theory of mind. According to Wechsler (2010), listeners already need first-order theory of mind to process pronouns in unembedded sentences like *I get the football*. In direct speech reports such as *Elephant said, "I get the football"*, listeners do not only have to shift from their own perspective to that of the speaker, but also from the actual speaker's perspective to the reported speaker's perspective. This double perspective shifting arguably requires second-order theory of mind. Previous studies indicate that children pass second-order false belief tasks between five and seven (Perner & Wimmer, 1985; Sullivan, Zaitchik, & Tager-Flusberg, 1994). However, it might take a few more years until children are able to apply second-order reasoning in the domain of language (cf. Flobbe, Verbrugge, Hendriks, & Krämer, 2008).

The third and, in our view, most important explanation is that the reporting speech context is so salient that it 'attracts' the pronouns in the report, yielding indirect-speech-like interpretations even for direct reports. When children heard, for instance, the first-person pronoun *I* in direct speech, they tended to incorrectly link it to the person who is currently speaking, that is the person who produces the speech report. Interestingly, a similar phenomenon has been observed in several languages. As discussed in section 2.3.5, various sign languages allow pronouns and other deictic elements under role shift – the sign language equivalent of direct speech – to get an unshifted, indirect speech interpretation. This leads precisely to cases where an *I* in an otherwise direct report picks out the reporting signer (Engberg-Pedersen, 1995). Evans (2012), who investigates direct speech cross-linguistically, draws attention to 'speech-act participant attraction' in Slave (Canada), Kwaza (Brazil), and Nez Perce (United States). In these languages, a second-person pronoun in an otherwise direct report can receive an unshifted interpretation, referring to the addressee of the reporting instead of the reported context (Evans, 2012).

These cross-linguistic findings of unshifting in (seemingly) direct speech correspond to how Dutch children interpret first-, second- and third-person pronouns embedded in Dutch direct speech. Following Evans's description and terminology, we speculate that the salience of the reporting context and its speaker and addressee causes it to attract the interpretation of pronouns, even in the presence of clearly detectable direct speech cues that should lead to a shifted interpretation. The need to inhibit this attraction might also explain why even Dutch adults have lower accuracy rates and longer decision times for direct speech items.

Assuming that the reporting context 'attracts' pronouns in our experiment, we propose that participants must have the following cognitive abilities to successfully overcome this attraction in the case of direct reports. First, they need to be able to inhibit the prepotent indirect speech interpretation. Second, they must have the ability to shift to the less salient reported context to determine the pronoun value. Third, they need the working memory skills required by a task, like ours, that involves strong incorrect prepotencies (Roberts & Pennington, 1996). All these aspects of executive function, inhibition, shifting and working

memory, have a protracted development and are not at an adult level yet at the age of eleven (cf. Brocki & Bohlin, 2004; Huizinga, Dolan, & Van der Molen, 2006). Support for this explanation comes from Epley, Morewedge and Keysar (2004) who tested children from a similar age range with a referential communication task. They conclude that children have more difficulties than adults with revising an initial incorrect interpretation. A similar explanation could also account for the high percentage of direct speech errors in our experiment. Both children and adults are at first drawn to the incorrect referent in the salient reporting context, but in contrast to adults, children are unable to correct their initial interpretation. This could be due to a serial processing bottleneck. Even though children might possess the cognitive skills necessary to perform the task, they are unable to serially apply the required mental operations (cf. Hendriks et al., 2007; Verbrugge, 2009).⁵

Until further studies are conducted, we assume that these three explanations complement each other. First, children seem to be less sensitive than adults to the prosodic cues that signal direct speech. Second, younger children might struggle with perspective differences between the reporting and reported speech context. And third, children might have insufficient executive functioning abilities to inhibit the attraction of the salient reporting context and to shift to the reported context.

Hypothesis 2: Local vs third-person pronouns

In line with previous studies (Brenner, 1983; Charney, 1980; Werner Deutsch & Pechmann, 1978; Legendre & Smolensky, 2012; Murphy, 1986), we found support for Lyons's (1977) split between local and third-person pronouns, but with some qualifications for pronouns embedded in direct and indirect speech reports.

In non-reportative statements (*He gets the football*), children did not only make significantly more mistakes for *hij* 'he' as compared to *ik* 'I' and *jij* 'you', they also needed more time to select a referent. This confirms the hypothesis that the interpretation of deictic third-person pronouns is based on a cognitively more demanding mechanism of reference assignment (Legendre & Smolensky, 2012; Lyons, 1977).

The fact that even adults were not at ceiling for the third-person in the 'no report' condition may be due to the lack of pointing or eye gaze to raise the salience of the intended referent of this deictic *hij* (Diessel, 2012; Levinson, 2004). The decision not to include an ostensive gesture is the result of a trade-off between naturalness of this 'no report' *hij* and the uniformity of the various items in the task. The unnaturalness of deictic *hij* without a pointing gesture presumably confused children and adults, leading to a deviant interpretation.

⁵ An interesting open question is whether children's comprehension of pronouns in direct speech could be improved by giving them feedback on whether their answers are correct or wrong (with possibly an explanation why). Previous studies show that children's first- and second-order false belief reasoning improves through training (Arslan & Verbrugge, 2015; Clements, Rustin, & McCallum, 2000).

A closer look at the incorrect results for the problematic ‘no report’ *hij* items shows that, in line with Murphy’s (1986) findings, participants selected the addressee instead of the ‘other person’ as referent for *hij* in 98% of the errors. The fact that older children and adults still made this particular type of error could be due to a misunderstanding of the speech situation. Some adults with this mistake told us after the experiment that they – at least initially – thought the animal was addressing them. Based on this assumption, they linked *hij* to the animal in the vicinity of the speaker, that is, the actual addressee. Note, however, that these potential task-related complications affect only the interpretation of *hij* in the ‘no report’ condition, but not in the crucial ‘speech report’ condition. The more extended sequence of events in the ‘speech report’ condition (A whispers something in B’s ear, B reports it to C) makes it unambiguously clear that the three animals are interacting with each other and not with the participant.

In indirect speech, children likewise made more mistakes for *hij* than for the local pronouns, with no significant difference between *ik* and *jij*. But compared to the ‘no report’ condition, where the mean percentage of correct interpretation of *hij* ranges between 2% (four-year-olds) and 33% (eleven-year-olds), accuracy is clearly higher in indirect speech with values between 85% (eleven-year-olds) and 95% (seven-year-olds). This means that third-person pronouns are easier to interpret in indirect speech reports (*Elephant said that he gets the football*) than in plain, non-embedded statements (*He gets the football*). The explanation for this *prima facie* surprising result is that while the non-embedded *hij* is used deictically, referring to an extra-linguistically salient third person, *hij* in indirect speech is used anaphorically, referring back to an intrasentential linguistic antecedent (*Elephant*). This explicit linguistic mention appears to make the correct referent of *hij* in indirect speech more salient and cognitively accessible. By contrast, the referent of deictic *hij* in an unembedded sentence needs to be linked to a non-participant in the extra-linguistic context. This seems to be especially demanding in our experimental scenario, where all three animals are possible referential candidates and additional ostensive gestures are absent.

Our results are consistent with Charney (1980) who found that children comprehend anaphoric third-person pronouns before deictic ones. Production studies show mixed results. Some researchers claim that children first produce third-person pronouns deictically and only later acquire their anaphoric use (Hickmann, 1995; Karmiloff-Smith, 1985). Others demonstrate that right from the onset of pronoun production children prefer to use third-person pronouns to refer to entities previously mentioned in the dialogue (Salazar Orvig et al., 2010). In any case, in retrospect it is not surprising that in our task the interpretation of a deictic *hij* (without an accompanying pointing gesture and three equally salient, gender-matched potential referents) is harder than the interpretation of an anaphoric *hij* (referring to the subject of the very sentence in which it occurs).

In direct speech, children made significantly fewer errors for first-person pronouns than for both second- and third-person pronouns. This is similar to what we found in adults.

Again, the apparent ease of *ik* in direct speech can be explained by the fact that the referent of first-person pronouns is more salient because it is also mentioned linguistically in the reporting clause (*Elephant said, "I get the football"*).

With regard to the difference between second-person *jij* and third-person *hij* in direct speech, we found that adults and eleven-year-old children made more mistakes with *hij* than with *jij*, as expected. However, in younger children, *jij* and *hij* in direct speech had either similar accuracy rates or the accuracy of *hij* even exceeded that of *jij*. We speculate that children's better performance on direct *he* as in *Monkey said, "He gets the football"* is due to a double confusion leading them to the correct interpretation 'by accident'. First, consider the error pattern in the 'no report' condition. In 98% of the mistakes children selected the addressee as referent of *hij*. Second, as described before, children tend to interpret direct speech as if it were indirect speech, that is, pronouns tend to be evaluated with respect to the reporting context. Combining these two systematic errors leads children to pick out the actual addressee, which happens to be correct due to the way our stimuli are designed. By way of illustration, consider a case where Monkey whispers to Dog and then Dog reports to Elephant, saying *Monkey said, "He gets the football"*. By the first systematic error, the child interprets *he* as referring to the addressee, and by the second, she ignores the context shift of direct speech. As a result of these two errors, she interprets the quoted *he* as referring to the addressee of the report, the elephant. This just happens to be the correct answer because Elephant is also the original third person in the reported context (i.e., Monkey whispering to Dog).

In sum, a refined picture emerges. While the split between first- and second- versus third-person pronouns is clearly observable in unembedded sentences, the situation is more complex in speech reports, where dependencies between reporting clause and report influence reference assignment.

4.2.6 Conclusion

The classical view in theoretical linguistics is that direct and indirect speech are two fundamentally distinct modes of reporting what someone said. Direct speech involves quotation, that is, reproducing an utterance from the original speaker's perspective, while indirect speech involves presenting what was said from one's own perspective. However, if we look at different languages and registers than standard, written Dutch or English we find many forms of reporting that do not fit in this strict dichotomy. There is some evidence that reported speech in child language is a case in point. For instance, we know that children allow *wh*-extraction (a known indicator of indirect speech) and indexical shifts (a known indicator of direct speech) within a single report.

On the basis of the abundance of such mixed reporting forms, Maier (2009) and Evans (2012) have proposed alternative accounts of speech reporting where direct and indirect speech are but extremes on a continuum of mixed reporting forms. Our experiment provides support for such a position by showing that children up to the age of twelve do not clearly dis-

tinguish direct and indirect speech in comprehension, even in a language like Dutch which has an exceptionally clear marking of the distinction in the adult grammar. More precisely, we found that children mixed direct and indirect speech in a very specific way, that is they interpreted pronouns in direct speech like in indirect speech, while errors in the opposite direction (interpreting pronouns in indirect speech like in direct speech) were rare. This indicates that in our task, an unshifted interpretation of pronouns, like in indirect speech and non-reportative sentences, is the default for children. Similar to speakers of languages such as Danish Sign language or Kwaza, Dutch children seem to ‘unquote’ pronouns in direct speech reports as they are pragmatically attracted to the more salient reporting context.

This suggests an acquisition path on which children start with a fluid form of reporting not clearly distinguishing between direct and indirect speech. In this phase, extra-linguistic factors such as the salience of the reporting speech context can drastically influence children’s pronoun interpretation. Children’s late acquisition of a strict direct–indirect distinction may be related to their development of executive function (to suppress attraction) and perspectival abilities (to perform the semantic context shift). Another factor deserving future research is the relation between literacy training and the acquisition of a rigid direct–indirect distinction. As Maier (2015b) points out, oral language seems to make do with a more fluid distinction between direct and indirect speech than written language. We speculate that children’s increased exposure to written language, where this distinction is more clearly marked and adhered to, could drive the development of an adult-like direct–indirect dichotomy.

5 EXPERIMENT 2: TESTING THE COMPREHENSION AND PRODUCTION OF REPORTED SPEECH IN INTERACTION

5.1 INTRODUCTION AND HYPOTHESES

A striking finding of experiment 1 is that even eleven-year-old children are not yet adult-like in their interpretation of pronouns in speech reports. This is surprising given that (i) direct and indirect speech have clearly distinct features in Dutch (see 2.2) and (ii) Dutch children are frequently exposed to speech reports in everyday conversations (see input analysis in section 3.1.5).

The question arises whether aspects of the experimental setup might be responsible for children's poor performance. In experiment 1, children observed the interactions of three animals on a tablet. They were not part of the interaction themselves, but overheard speech reports that were addressed to someone else. Even though it is not uncommon for children to witness conversations of other people, for instance between their parents or a parent and a sibling, it seems to be easier for children to process deictic expressions when they assume the role of the addressee. Murphy (1986) and Charney (1980) found that children made significantly fewer errors in interpreting deictic pronouns when they were addressed as opposed to just overhearing speech. In contrast to experiment 1, participants in experiment 2 are therefore part of the interaction and the speech reports that they have to interpret are directly addressed to them. The second change we made is that we transferred the interaction from the virtual world (presented on a tablet) to the physical world. Following research on embodied cognition (Barsalou, 1999, 2008; Wilson, 2002), we created a more naturalistic situation that contains three-dimensional objects as props and hand puppets as interaction partners. We predict that it is easier for children to take the perspective of someone else when he or she is physically present in the situation. In particular, the bodily presence of the reported speaker might increase his salience and facilitate a shift to his perspective which is required for interpreting pronouns in direct speech. We hypothesize that these two modifications – involvement in the interaction and a more naturalistic environment – will improve children's pronoun interpretation in direct speech.

In addition to the comprehension part, experiment 2 also includes a production part. The corpus study in chapter 3 has revealed that Dutch and German children produce more direct than indirect speech in natural interactions and that they rarely mix the two types of

reports. In experiment 2, we now test Dutch children's production of speech reports in an experimentally controlled situation. This has the advantage that all speakers produce speech reports under the exact same conditions (e.g., same discourse context, same interaction partners, same original utterance). This allows us to better compare the types of speech reports that participants produce at different ages. Furthermore, if Dutch children actually do produce direct-indirect mixes, we should be able to detect them with this methodology because the speech reports can be directly contrasted with the original utterance.

Similar to the comprehension part, the function of a speech report in the production part is to transmit a propositional content (who gets a certain object). It is evident from experiment 1 that in this type of discourse context both adults and children find it easier to interpret pronouns in indirect speech. Therefore, if speakers take the addressee's perspective into account, we would expect them to produce mainly indirect speech reports in this situation. However, from the perspective of the reporting speaker, direct speech seems to be less demanding than indirect speech. In direct speech the original message, including the pronouns, can simply be copied, while in indirect speech a deictic re-orientation from the reported speaker's perspective to the reporting speaker's perspective is necessary. This requires deictic adaptations, such as replacing *you* by *I*.

Taken together, in the production of speech reports in the communicative situation of experiment 2, two potentially conflicting factors are involved: (i) audience design, which would suggest the use of indirect speech, and (ii) difficulty for the reporting speaker, which would suggest the use of direct speech. Our hypothesis is that adults will mainly produce addressee-friendly indirect speech reports, younger children will prefer the cognitively less demanding direct speech reports, and older children will be somewhere in between younger children and adults.

Despite the fact that interpreting speech reports and producing speech reports are two different skills, the connecting factor seems to be that in both cases knowledge about the features of direct and indirect speech are required. With respect to experiment 2, we therefore predict that children's performance in the comprehension and production part is related. Children who make fewer mistakes in pronoun interpretation in direct and indirect speech should also produce more addressee-oriented indirect speech reports, and vice versa. We test this by including production as a factor in our comprehension model and comprehension as a factor in our production model.

The participants in experiment 2 are divided in two groups. The first group is tested with the comprehension part before the production part, the other group with the production part before the comprehension part. We hypothesize that the test order influences participants' performance in two specific ways. If participants are listeners first, it should be easier for them to take the listener's perspective into account when producing speech reports. We therefore predict that children who are tested with a preceding comprehension part produce more addressee-friendly indirect speech reports than their peers starting with the production part. In turn, if children experience the production part first, they should be

better than their peers in interpreting direct and indirect speech reports in the comprehension part. The reason for this is that in the production part, the participants take part in both the reported and the reporting speech context. This could raise children's awareness that one and the same person can have different communicative roles (speaker, addressee, other) in different contexts. We test these predictions by including test order in our comprehension and production models.

5.2 METHODS

5.2.1 Participants

The participants of this study were 38 five-year-old children, 38 nine-year-old children and 19 adults (see table 1). All participants were native speakers of Dutch. Eighteen additional children were tested, but had to be excluded due to cognitive and linguistic problems (5), experimenter error (1) or inability to understand the experimental task (12). The children who struggled with the task were exclusively younger children. This indicates that we have reached the lower age limit for testing children with this task. Children were recruited from an elementary school in the North of the Netherlands and received a small reward (a sticker for five-year-olds, a pen for nine-year-olds) for participating. Written parental consent was obtained prior to the experiment. Adults – mostly students from the University of Groningen – received monetary compensation for their participation. All participants were tested individually in a quiet room at the school or university by three female experimenters.

Table 1: Participants

Age group	Number	Mean age	Range	Gender (f/m)	Test order (comp./prod.)
5	38	5.4	4;5–6;8	19/19	18/20
9	38	9.0	8;4–10;2	20/18	19/19
adults	19	23.8	19–38	15/4	9/10

5.2.2 Stimuli and Procedure

Experiment 2 consists of three parts: an introduction part, a comprehension part and a production part. All participants were randomly assigned to one of two groups: About half of the participants saw the comprehension items before the production items, the other half was tested in the reverse order (see table 1). In both parts, the participants interacted with two hand puppets played by two experimenters; a third experimenter gave instructions. Female participants interacted with the female puppets Mimi and Lola, male participants with the

male puppets Bobo and Rudi (see fig. 1). The reason for using gender-matched puppets is to ensure that the gender-information on the third-person singular pronouns *zij* ‘she’ and *hij* ‘he’ did not serve as an additional cue for reference resolution.



Figure 1: Hand puppets used in experiment 2 (from left to right: Mimi, Lola, Bobo, Rudi)

Introduction phase

In the introduction phase, the instructing experimenter introduced the two puppets (Mimi and Lola or Bobo and Rudi) and ensured that the participant knew the puppets’ names and gender. Subsequently, the experimenter pointed out three cardboard boxes, one for each puppet and one for the participant him- or herself. The boxes of the puppets had a picture of their face on it and the participant’s box was marked with a yellow star (see fig. 2). The order in which the boxes were arranged was counterbalanced across participants. The experimenter also drew the participant’s attention to toy objects that were displayed on a table in front of the boxes (fig. 3).

Similar to experiment 1, the experimental game is about determining who gets a certain object. However, while the participants of experiment 1 watched the interaction of three animals as external observers, the participants of experiment 2 are part of the interaction and have active roles in the game: In the comprehension part, they are the addressees of the speech report and have to put the objects into the correct boxes; in the production part, they have to transmit messages from one puppet to the other. As a consequence of their direct involvement, participants are potential referents of pronouns and can receive objects.



Figure 2: Boxes (version for female participants)



(a)

(b)

Figure 3: Objects used in comprehension part (a) and production part (b)

Comprehension part

In the comprehension part, the participant is positioned behind the three cardboard boxes and is asked to put the objects into the correct box, either his or her own box or that of one of the puppets. The rules of the game are that Lola¹ decides who gets which object. However, Lola and the participant are not allowed to communicate directly with each other. This is why Lola whispers – unintelligibly for the participant – into Mimi’s ear who gets a certain object. Subsequently, Mimi picks up the object and transmits the message to the participant using either a direct (1) or indirect (2) speech report.

- (1) Lola zei, “Ik krijg de koe”.
‘Lola said, “I get the cow”.
- (2) Lola zei dat jij de tafel krijgt.
‘Lola said that you get the table.’

On the basis of this report, the participant has to identify the recipient of the object and put the object into the respective box. In total, 30 speech reports, corresponding to 30 objects, were presented. Half of the reports were in direct speech, half in indirect speech. We showed 5 test items for each combination of report type (direct speech, indirect speech) and pronoun

¹ We explain the experiment on the basis of the version for female participants. In the version for male participants, Lola and Mimi are replaced by Bobo and Rudi. Additionally, the masculine pronoun *hij* ‘he’ is used instead of the feminine pronoun *zij* ‘she’.

type (*ik* 'I', *jij* 'you' and *hij/zij* 'he/she'). Note that the direct and indirect speech items were not presented in blocks, but appeared in a fixed random order. In total, we had 4 different lists, which differed in the order of the test items and the gender information. All speech reports were produced by the same female experimenter who played the puppet Mimi or Rudi. This experimenter had been trained beforehand to produce consistent speech reports. The direct speech reports contain a break before the quotation and a change of voice to a higher pitch within the quotation. No difference was made for voicing female or male puppets.²

Production part

In the production part, the game remains the same, but the participant and puppets change roles. The participant now assumes the role of the person who transmits information from one person to another. Puppet Mimi tells the participant who gets a certain object with a simple non-embedded sentence such as (3). The 15 test sentences in the production part contain either a first-, second- or third-person singular pronoun (5 per pronoun *ik* 'I', *jij* 'you' and *hij/zij* 'he/she').

- (3) Jij krijgt de piraat.
'You get the pirate.'

The second puppet Lola is unable to hear Mimi's original utterance because she is wearing headphones. The participant needs to pick up the correct object (e.g., the pirate) and tap Lola on the shoulder, so that she takes off her headphones. Subsequently Lola asks the participant *Zei Mimi iets?* 'Did Mimi say something?'. We chose this question with the purpose of eliciting preferably full-fledged speech reports including a reporting clause. If a participant only answered *yes* without reporting the relevant utterance, Lola inquired *Wat zei Mimi dan?* 'What did Mimi say?'. Participants were told that they should make it as easy as possible for Lola to understand who gets the objects, with the restriction that they were only allowed to use words and no pointing gestures. Note that the participants could not see into which box puppet Lola eventually put the object, i.e., they did not receive feedback about how Lola interpreted their report. The participants' verbal responses were audio-taped and transcribed.³

5.3 RESULTS

We first present the results from the comprehension part, which we subsequently compare with the performance of age-matched participants from experiment 1. Afterwards we analyze participants' production of speech reports.

² See Appendix B: Experiment 2 for an overview of the test sentences in the comprehension part.

³ See Appendix B: Experiment 2 for an overview of the test sentences in the production part.

5.3.1 Comprehension

5.3.1.1 Accuracy of pronoun interpretation

Figures 4 and 5 show the accuracy of pronoun interpretation of five- and nine-year-old children and adults in direct speech (fig. 4) and indirect speech (fig. 5).

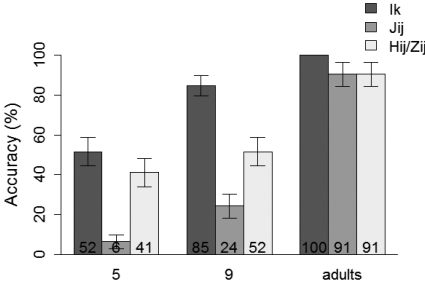


Figure 4: Percentage of correct pronoun interpretation in direct speech per age group. Error bars indicate 95% confidence intervals.

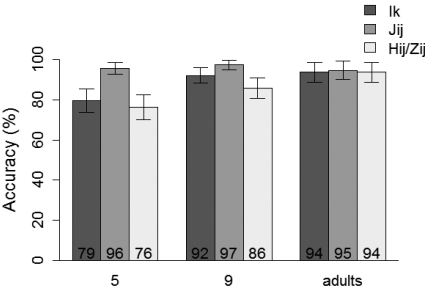


Figure 5: Percentage of correct pronoun interpretation in indirect speech per age group. Error bars indicate 95% confidence intervals.

Table 2: Goodness of fit of the fixed-effect factors of the model fitted to children’s accuracy of pronoun comprehension in speech reports. An Akaike Information Criterion (AIC) decrease of more than 2 indicates that the goodness of fit of the model improves significantly (Akaike, 1974).

	AIC decrease
Random-effect factors (random intercepts for subjects, by-subject random slopes per REPORT TYPE)	
+ REPORT TYPE	69.9
+ PRONOUN TYPE	125.5
+ REPORT TYPE * PRONOUN TYPE	255.9
+ AGE	22.9
+ EXPERIENCE	2.7
+ ADDRESSEE-ORIENTED PRODUCTION	4.2

We analyzed children’s accuracy of pronoun interpretation with mixed-effects logistic regression modeling. With a procedure of model comparison, we added stepwise fixed-effect factors to the baseline model (containing random intercepts for subjects and by-subject random slopes per REPORT TYPE) and checked whether they improve the goodness of fit of the model (criterion: AIC decrease of more than 2). Table 2 shows that the factors REPORT TYPE (direct, indirect), PRONOUN TYPE (first-, second-, third-person), the interaction between REPORT TYPE and PRONOUN TYPE, AGE, EXPERIENCE (1–5) and ADDRESSEE-ORIENTED PRODUCTION predict children’s accuracy of pronoun interpretation in speech reports. The factor EXPERIENCE indicates how often a participant has seen a certain combination of report type and pronoun type (e.g., direct *I*) before. ADDRESSEE-ORIENTED PRODUCTION specifies how many of a participant’s reports in the production part were deictically oriented towards the reporting speech context. This is the case for all indirect speech reports.⁴ The index of concordance of the model is 0.93.

Age has a positive effect on accuracy, meaning that older children tended to perform significantly better than younger ones ($\beta = 0.35$, $z = 5.29$, $p < .001$). Children’s performance improved during the experiment. The more experience they had with a certain type of test item, the higher the chance that they interpreted it correctly ($\beta = 0.10$, $z = 2.17$, $p = .030$). Children’s behavior in the production part predicted their performance in comprehension. Children who produced more addressee-oriented reports were also more likely to be better in comprehension ($\beta = 0.05$, $z = 2.54$, $p = .011$). By contrast, the order in which participants saw the production and comprehension part, the gender of the participants and the order of the cardboard boxes did not predict accuracy of pronoun interpretation.

⁴ See section 5.3.2.3 for a more detailed description of addressee-oriented reports.

5.3.1.2 Direct–indirect mixes

As pointed out for experiment 1, we define direct–indirect mixes in our comprehension study as systematic mistakes in the interpretation of pronouns. In direct speech, a direct–indirect mix means that pronouns are evaluated with respect to the reporting speech context, like in indirect speech. In indirect speech, the opposite error counts as a direct–indirect mix, i.e., if the reported speech context is incorrectly chosen as context of evaluation, like in direct speech. Our analysis shows that 92.0% ($N = 647$) of the mistakes in direct speech and 81.3% ($N = 139$) of the mistakes in indirect speech are direct–indirect mixes. A one sample *t*-test indicates that in both direct and indirect speech, the distribution of mistakes differed significantly from the chance level of 50% (direct speech: $t(646) = 39.23$, $p < .001$; indirect speech: $t(138) = 9.43$, $p < .001$). This means that in contrast to experiment 1, where the mistakes in indirect speech were randomly distributed, we find the systematic error pattern of direct–indirect mixing now in both direct and indirect speech. The number of direct–indirect mixes decreased with age. While five-year-old children systematically misinterpreted 37.6% of the reports, the percentage of direct–indirect mixes in comprehension decreased to 24.5% in nine-year-old children and 5.4% in adults.

5.3.1.3 Comparison with experiment 1

We compared participants' comprehension performance in experiment 2 with that of same-aged participants from experiment 1 (age groups: 5, 9 and adults (tested with tablet)). For each age group, we pooled the comprehension data from experiment 1 and 2 and created a separate logistic regression model. This way, we can find out whether the type of experiment (experiment 1, experiment 2) influences the accuracy of pronoun interpretation in five-year-olds, nine-year-olds and adults.

Table 3: AIC decrease for the inclusion of fixed-effect factors and interactions to the models explaining accuracy of pronoun interpretation in experiments 1 and 2 in different age groups. An Akaike Information Criterion (AIC) decrease of more than 2 indicates that the goodness of fit of the model improves significantly (Akaike, 1974).

	5-year-olds	9-year-olds	Adults
Random-effect factors (random intercepts and slopes per REPORT TYPE for subjects)			
+ REPORT TYPE	67.9	45.6	12.3
+ PRONOUN TYPE	25.4	153.1	44.4
+ REPORT TYPE * PRONOUN TYPE	152.7	87.2	8.4
+ EXPERIMENT	-1.1	4.6	-0.9
+ REPORT TYPE * PRONOUN TYPE * EXPERIMENT	22.8	13.8	1.9
+ EXPERIENCE	-0.9	1.3	15.0

Table 3 shows which fixed-effect factors predict the comprehension of speech reports in the three different age groups. For both five- and nine-year-old children, we found a three-way interaction between REPORT TYPE, PRONOUN TYPE and EXPERIMENT. Adults' performance did not differ significantly in experiment 1 and 2, as evident from the fact that EXPERIMENT is not a fixed-effect factor in adults' accuracy model. However in contrast to children, adults' accuracy of pronoun interpretation improved with experience, $\beta = 0.39$, $z = 3.97$, $p < .001$.

The three-way interaction in children's accuracy models tells us that the type of experiment has an impact on children's pronoun interpretation in speech reports. In order to find out in what respects children's performance in experiments 1 and 2 differs, we conducted a follow-up analysis. With multiple comparisons from the 'multcomp' package, we compared the means of all combinations of REPORT TYPE, PRONOUN TYPE and EXPERIMENT in the data of the five- and nine-year-olds (Hothorn et al., 2008).

In the following, we report only those results that are significant and that either (a) compare the same report type in experiment 1 and experiment 2, or (b) compare within experiments 1 or 2 the accuracy of first-, second- and third-person pronouns in direct or indirect speech.

Comparison with experiment 1: Five-year-olds

Table 4 and figure 6 present the relevant significant differences for the five-year-olds. If we compare the performance of five-year-old children in experiments 1 and 2, we can observe two main differences. First, children in experiment 2 outperformed their peers in experiment 1 in the interpretation of direct speech reports with the pronouns *I* and *he/she*. However, no improvement was found for the second-person pronoun *you* in direct speech. Second, five-year-old children made significantly more mistakes in interpreting indirect *I* in experiment 2 than in experiment 1.

Table 4: Multiple comparisons of means (Tukey contrasts) for accuracy of pronoun interpretation in experiments 1 and 2 for the five-year-olds.

	Estimate	SE	z value	p value
Experiment 1 vs. Experiment 2				
Direct <i>I</i>	2.21	0.36	6.12	<.001
Direct <i>he/she</i>	1.59	0.35	4.57	<.001
Indirect <i>I</i>	-2.28	0.63	-3.61	.014
Experiment 1				
No significant differences				
Experiment 2				
Direct <i>I</i> – Direct <i>you</i>	-2.89	0.34	-8.50	<.001
Direct <i>he/she</i> – Direct <i>you</i>	-2.44	0.34	-7.17	<.001
Indirect <i>I</i> – Indirect <i>you</i>	1.82	0.41	4.47	<.001
Indirect <i>he/she</i> – Indirect <i>you</i>	2.01	0.40	4.99	<.001

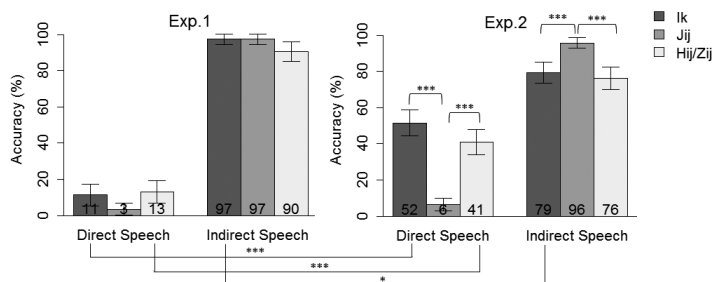


Figure 6: Percentage of correct pronoun interpretation of five-year-olds in experiment 1 (left) and experiment 2 (right); relevant significant differences marked at different significance levels: * $p < .05$, ** $p < .01$, *** $p < .001$. Error bars indicate 95% confidence intervals.

While in experiment 1 the accuracy of first-, second- and third-person pronouns in direct or indirect speech did not differ, we found significant differences between pronouns in experiment 2. In particular, second-person *you* was the pronoun with the lowest accuracy in direct speech, significantly lower than that of *I* and *he/she*. In indirect speech by contrast, children were significantly better in interpreting *you* than both *I* and *he/she*.

Comparison with experiment 1: nine-year-olds

Table 5 and figure 7 show relevant significant differences between experiments 1 and 2 for nine-year-old children. Similar to the five-year-olds, nine-year-olds in experiment 2 outperformed their peers in experiment 1 with direct speech reports including first- and third-person pronouns but not second-person pronouns. In contrast to the younger children, nine-year-olds' indirect speech performance did not decrease significantly in experiment 2.

Table 5: Multiple comparisons of means (Tukey contrasts) for accuracy of pronoun interpretation in experiments 1 and 2 for the nine-year-olds.

	Estimate	SE	z value	p value
Experiment 1 vs. Experiment 2				
Direct <i>I</i>	2.35	0.39	6.03	<.001
Direct <i>he/she</i>	1.43	0.39	3.70	.010
Experiment 1				
Direct <i>I</i> – Direct <i>you</i>	-1.48	0.37	-4.04	<.01
Experiment 2				
Direct <i>I</i> – Direct <i>you</i>	-3.30	0.29	-11.25	<.001
Direct <i>I</i> – Direct <i>he/she</i>	-1.89	0.27	-7.06	<.001
Direct <i>he/she</i> – Direct <i>you</i>	-1.41	0.25	-5.76	<.001
Indirect <i>he/she</i> – Indirect <i>you</i>	1.90	0.50	3.77	<.01

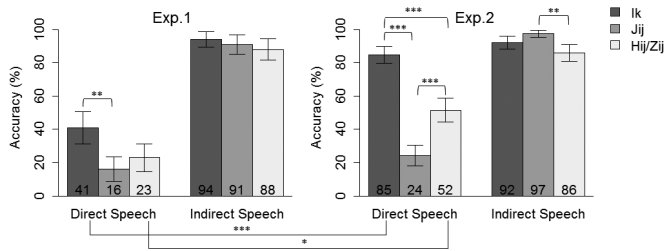


Figure 7: Percentage of correct pronoun interpretation of nine-year-olds in experiment 1 (left) and experiment 2 (right); relevant significant differences marked at different significance levels: * $p < .05$, ** $p < .01$, *** $p < .001$. Error bars indicate 95% confidence intervals.

In experiment 1, only direct *I* and direct *you* differed significantly. In experiment 2, *you* is the pronoun with the lowest accuracy in direct speech, significantly lower than that of direct *I* and direct *he/she*. First-person *I* has clearly the highest accuracy in direct speech, significantly higher than that of *you* and *he/she*. In indirect speech, nine-year-olds were better in interpreting second- than third-person pronouns.

In sum five- and nine-year-old children’s pronoun interpretation in direct speech was significantly better in experiment 2 as compared to experiment 1. However, this only applies to first- and third-person pronouns, but not to second-person pronouns.

5.3.2 Production

5.3.2.1 Overview of produced speech reports

After the analysis of participants’ comprehension of speech reports, let us turn to their production data. We classified all speech reports that participants produced into one of the six categories, presented in table 6.

Table 6: Categories of speech reports for participants' production data. Original utterance: *Jij krijgt de krokodil* 'You get the crocodile'.

Report type	Example
Direct (framed)	Mimi zei, "Jij krijgt de krokodil." 'Mimi said, "You get the crocodile."'
Direct (unframed)	"Jij krijgt de krokodil." "You get the crocodile."
Indirect (framed)	(Mimi zei) dat ik de krokodil krijg. '(Mimi said) that I get the crocodile.'
Indirect (description)	Ik krijg de krokodil. 'I get the crocodile.'
Mix	Mimi zei dat jij de krokodil krijgt. 'Mimi said that you get the crocodile.' Mimi zei, "Ik krijg de krokodil." 'Mimi said, "I get the crocodile."'
Other	Mimi krijgt de krokodil. 'Mimi gets the crocodile.'

We distinguish two types of direct speech (direct (framed) and direct (unframed)) and two types of indirect speech (indirect (framed) and indirect (description)). The category 'mix' includes all cases in which participants combine features of direct and indirect speech in an unconventional way. For instance, if they use the syntax of indirect speech, but do not relate pronouns to the reporting speech context. The category 'other' includes all reports that do not fit into one of the other five categories.⁵

Table 7: Percentage of speech report types per age group, distinguished by test order (prod = production first, comp = comprehension first) (N = 1425)

	5-year-olds		9-year-olds		Adults	
	prod	comp	prod	comp	prod	comp
Direct (framed)	0	1.9	0	14.0	0	0
direct (unframed)	39.3	31.1	30.2	31.2	0	0
Indirect (framed)	8.7	10.7	0.4	11.6	12.7	82.2
indirect (description)	43.0	56.3	67.7	41.4	84.7	16.3
Mix	2.0	0	0	0.4	0.7	0.7
Other	7.0	0	1.8	1.4	2.0	0.7

⁵ Note that participants did not always use pronouns to refer to the recipient of the object. In 18.4% of the reports, they used names (e.g., *Mimi krijgt de helm* 'Mimi gets the helmet').

Table 7 shows the production of different types of speech reports in five- and nine-year-old children and adults, presented separately for both test orders (production first, comprehension first). Adults' speech reports show a clear pattern: They produced either indirect speech with a reporting clause or an indirect description. Their preference for one or the other was influenced by the test order. Adults who started with the production part preferred unframed descriptions (84.7%). Adults who were exposed to framed speech reports in the comprehension part prior to production used more framed indirect speech reports themselves (82.2%).

A clear difference between adults and children is that children produced more than 30% unframed direct speech reports. This report type is highly ambiguous. An unframed quotation like "*You get the crocodile*" looks syntactically like a description, but copies the pronouns from the original speech context. Since children who used unframed direct speech did not mark the perspective shift prosodically, this report type is likely to be misinterpreted by the addressee who in the absence of any quotation signals will probably evaluate the pronoun with respect to the actual reporting context. The data suggest that five-year-olds who had the comprehension part first produced fewer ambiguous unframed direct reports (31.1% as compared to 39.3%).

Similar to adults, children who were tested with a preceding comprehension part tended to produce more framed reports than their peers who started with the production part. In nine-year-olds, for instance, the rate of framed direct speech went up from 0% to 14% and the rate of framed indirect speech from 0.4% to 11.6% when children were exposed to the comprehension items first.

5.3.2.2 Direct–indirect mixes

Participants produced not many direct–indirect mixes. We found only 9 instances in total (0.6% of all reports), of which 5 are accounted for by one child aged 4;11. Of these 9 mixes, 6 consisted of an indirect speech construction with a pronoun copied from the original utterance, like in example (4).

- (4) Original: Jij krijgt het skateboard
 'You get the skateboard'
 Report: Dat jij de skateboard krijgt
 'That you get the skateboard' (4;11)

Interestingly, one child and one adult produced a so-called mixed quotation (see (5) and (6)), in which the quoted pronoun is used and mentioned at the same time (Davidson, 1979). But only the adult resolved the quoted pronoun subsequently for her addressee by adjusting it to the shared reporting context (*I think that you get the ladder*).

- (5) Original: Hij krijgt de hond.
 ‘He gets the dog.’
 Report: Dat voor “hij” deze hond is.
 ‘That this dog is for “he”’ (5;3)
- (6) Original: Zij krijgt de ladder.
 ‘She gets the ladder.’
 Report: Ik heb hier de ladder en ik heb van Mimi gehoord dat “zij” de ladder krijgt. Maar volgens mij hebben we daar een probleem. Want ehm maar ik moet bepalen. Ik denk dat jij de ladder krijgt.
 ‘I have the ladder here and I have heard from Mimi that “she” gets the ladder. But in my view we have a problem here. For um but I have to decide. I think that you get the ladder.’ (adult)

The remaining mix, (7), combines the syntax of direct speech with a pronoun adjusted to the reporting context. Alternatively, (7) could also be an indirect speech report with an embedded main clause.

- (7) Original: Jij krijgt de krokodil.
 ‘You get the crocodile.’
 Report: Mimi zei, ik krijg de krokodil.
 ‘Mimi said, I get the crocodile (8;6)

5.3.2.3 Addressee-oriented reports

Participants were told to make the message as easy as possible for their addressee to understand. From our two comprehension experiments, we know that pronouns are easier to interpret in indirect than in direct speech in this kind of information exchange context. Consequently, an optimally addressee-designed speech report contains pronouns that are oriented towards the actual reporting speech context, i.e., the context in which the participant reports Mimi’s utterance to Lola.

We classified all speech reports participants produced with respect to whether or not they take the reporting context as reference point for pronouns. This includes all reports classified as ‘indirect (framed)’ and ‘indirect (description)’ as well as some reports in the categories ‘mix’ and ‘other’.

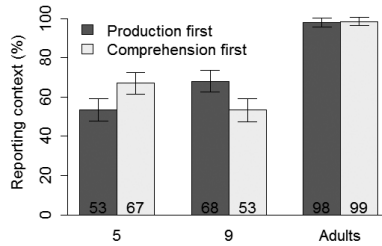


Figure 8: Percentage of produced reports oriented towards reporting context per age group. Error bars indicate 95% confidence intervals.

Figure 8 shows the percentage of reports per age group that are oriented towards the reporting speech context. It distinguishes between two groups of participants, those who saw the production part first and those who saw the comprehension part before the production part.

We used mixed-effects logistic regression modeling to find out which factors predict the choice of the reporting context as reference point. We were particularly interested in whether AGE GROUP (5, 9, adults), TEST ORDER (production first, comprehension first) and the ACCURACY SCORE from the comprehension part (0–30) have an effect on the use of addressee-oriented reports.

Table 8: Fixed-effect factors of the model explaining production of addressee-oriented reports

	Estimate	SE	z value	p value
(Intercept)	7.82	1.25	6.24	$p < .001$
5-year-olds vs. adults	-6.66	1.50	-4.44	$p < .001$
9-year-olds vs. adults	-6.16	1.50	-4.08	$p < .001$

Note: Model includes random intercepts for subjects.

As table 8 shows, only AGE GROUP turned out to be a predictor of the choice of speech context in production. Compared to adults, five- and nine-year-old children produced significantly fewer reports that were oriented towards the reporting context, with no difference between five- and nine-year-olds. The index of concordance C of the model is 0.98. Inclusion of the factors TEST ORDER, ACCURACY SCORE, SEQUENCE NUMBER, PRONOUN and GENDER of the participants did not improve the goodness of fit of the model.

5.4 DISCUSSION

In experiment 2, we looked at three different aspects of speech report acquisition: comprehension, production and the relation between comprehension and production. For each of these three aspects, we have formulated specific hypotheses. In the following, we discuss whether these predictions were confirmed by the results and try to explain some unexpected patterns in the data.

5.4.1 Comprehension

As outlined in the introduction (5.1), there are two principal differences in design between the comprehension tests used in experiments 1 and 2. First, participants in experiment 2 interpreted speech reports as addressees instead of as external observers. Second, the setting of experiment 2 was more naturalistic involving real objects and puppets as interaction partners. We predicted that these changes would make it easier for children to shift to the perspective of the reported speaker, which is necessary for the interpretation of pronouns in direct speech. This hypothesis is confirmed by the data. Compared to their age-matched peers from experiment 1, five- and nine-year-old children who participated in experiment 2 made significantly fewer mistakes in interpreting the pronouns *ik* 'I' and *hij/zij* 'he/she' – but not *jij* 'you' – in direct speech.

Both factors, the role in the interaction and the more naturalistic setup, could have brought about this improvement in direct speech interpretation. As previous studies show, interpreting deictic expressions is easier for children when they are addressees rather than non-addressed listeners (Charney, 1980; Murphy, 1986). Following this suggestion, the shift to the role of an addressee might also facilitate the interpretation of deictic pronouns that are embedded in direct speech reports.

A second possibility is that the bodily presence of the reported speaker serves as a physical anchor in the situation. This could make the reported speaker a more salient deictic orientation point and decrease the cognitive load of a shift to his perspective in direct speech.

In order to determine the exact contribution of each of these factors further experimentation is necessary. There are at least two more constellations in which children's comprehension of speech reports could be tested: a tablet experiment in which participants are directly addressed, and an experiment with puppets in which participants are external observers.

Interestingly, for the five-year-olds the improvement in direct speech was coupled with a decrease in the accuracy for indirect speech reports. While five-year-olds in experiment 1 interpreted 97% of the indirect *I* items correctly, their peers in experiment 2 had an accuracy rate of only 79%. In addition five-year-olds' accuracy for indirect *he/she* dropped from 90% in experiment 1 to 76% in experiment 2. This pattern could be explained by a shift away from the bias to evaluate pronouns with respect to the reporting speech context regard-

less of report type. The children in experiment 2 seem to increasingly consider the reported context as possible context of evaluation. Evidence for this is the fact that we find systematic direct–indirect mixes not only in direct but also in indirect speech comprehension. This means that – in contrast to experiment 1 – participants’ errors in indirect speech were not random but consisted in 81% of the cases in incorrectly choosing the reported speech context as context of evaluation.

Taken together, the higher salience of the reported context in experiment 2 seems to facilitate pronoun interpretation in direct speech. At the same time, it creates a competing context of evaluation in indirect speech. Especially in younger children, with a presumably less rigid direct–indirect distinction, this leads to a drop in the accuracy for pronoun interpretation in indirect speech.

Unlike five- and nine-year-old children, adults’ interpretation of direct and indirect speech reports was not influenced by the specific situational circumstances in which the direct and indirect speech stimuli were presented.

5.4.2 Excursus: Special properties of ‘you’

While children’s interpretation of first- and third-person pronouns in direct speech improved significantly in experiment 2, no such effect was found for second-person pronouns. Also in other respects second-person *jij* ‘you’ appears to be special. For both five- and nine-year-old children, *you* was the pronoun with the highest accuracy in indirect speech, but the lowest accuracy in direct speech. In this section, we discuss why the second-person pronoun stands out in experiment 2. For this purpose, we introduce Wechsler’s (2010) *de se* theory of person indexicals and check whether its empirical predictions are in line with our findings.

In an influential paper, Wechsler (2010) argues that pronouns that refer to the self are interpreted via a fundamentally different mechanism than pronouns that refer to other people. In the philosophical literature, the special status of the first-person pronoun *I* has long been acknowledged (e.g., Kaplan, 1989; Lewis, 1979; Perry, 1979). Perry (1979), for instance, illustrated it with the following story:

I once followed a trail of sugar on a supermarket floor, pushing my cart down the aisle on one side of a tall counter and back the aisle on the other, seeking the shopper with the torn sack to tell him he was making a mess. With each trip around the counter, the trail became thicker. But I seemed unable to catch up. Finally it dawned on me. I was the shopper I was trying to catch. (Perry, 1979, p. 3)

In order to express his new insight that he is making a mess, John Perry needs to use a first-person pronoun (*I am making a mess*) because other referring expressions such as third-person pronouns (*he*), names (*John Perry*) or descriptions (*the shopper with the torn sack*) do not unambiguously capture his new *de se* belief.

While the philosophical debate mostly focuses on first-person pronouns, Wechsler (2010) stresses that a similar case can be made for second-person pronouns:

Most work on self-ascription has focused on the first person, but second-person pronouns have exactly the same self-ascriptive force, only applied to the ADDRESSEE instead of the speaker. (Wechsler, 2010, p. 342)

Transferred to the sugar trail example, this implies that the most successful way of communicating to John Perry that he is making a mess would be to use a second-person pronoun (*You are making a mess*). The use of other referential expressions does not necessarily induce John Perry to form the *de se* belief that he messed up.

Looking at semantics from a communication-theoretic perspective, i.e., as coordinating the attitudes of speaker and hearer, *you* and *I* are each other's mirror image: while *I* corresponds to self-ascription in the speaker, *you* corresponds to self-ascription in the addressee. In all other cases (e.g., processing of *I* from the perspective of the addressee, processing of *you* from the perspective of the speaker) the pronoun's reference is determined indirectly, through reasoning about the other speech-act participant's attitudes. According to Wechsler, this is a cognitively more demanding process that requires theory of mind.

As empirical support for his theory, Wechsler refers to two populations: typically developing children before the age of four and children with autism. Wechsler's *de se* theory predicts that children acquire self-reference before other-reference. This is supported by studies showing that children produce *I* before they comprehend it, and comprehend *you* before they are able to produce it (Charney, 1980; Loveland, 1984). Since according to Wechsler's theory the production and comprehension of pronouns that refer to other people require a theory of mind, the finding that children with autism have problems in both theory of mind and pronouns suggests that these two skills are related (Baron-Cohen et al., 1985; Lee et al., 1994; Tager-Flusberg, 1994).⁶

⁶ There are a number of problems in using these empirical findings as support for Wechsler's theory. First, it remains rather mysterious that children produce first-person pronouns *ex nihilo* without understanding that they can refer to other selves as well. An alternative explanation for the prior production of *I* is that children produce it not as independent linguistic unit but as part of unanalyzed chunks such as *I-wanna* (Merleau-Ponty, 1964; Stawarska, 2009). Second, a direct relationship between pronoun processing and theory of mind in children with autism still needs to be empirically demonstrated, for instance with correlational or training studies. Third, there is evidence that individuals with autism process pronouns differently than non-autistic individuals (Mizuno et al., 2011). So even if theory of mind is a prerequisite for pronoun processing in children with autism, this does not allow straightforward conclusions about typically developing children.

On top of Wechsler's evidence from children, recent psycholinguistic studies with adults confirm the idea that *you* has special properties when it is interpreted from the perspective of the addressee. In the studies of Brunyé, Ditman, Mahoney, Augustyn, and Taylor (2009) and Sato and Bergen (2013), participants listened to statements such as *You are slicing the tomato* and had to verify whether a picture matches the described action. They found that with the second-person pronoun *you*, participants were quicker in verifying pictures when the slicing action was depicted from an internal first-person perspective than from an external perspective. This suggests that people mentally simulate statements with second-person pronouns from their own perspective. Moreover, when directly addressed as protagonist in a narrative, participants had a better spatial representation of the scenes, a better memory of the actions, and they showed more emotional responses and a higher identification with the characters (Andeweg, Hendrix, Van't Hoff, & De Hoop, 2013; Brunyé, Ditman, Mahoney, & Taylor, 2011; Ditman, Brunyé, Mahoney, & Taylor, 2010).

Against the background of this literature on the *de se* properties of *you*, let us revisit our experimental findings on the comprehension of second-person pronouns in speech reports. First, experiments 1 and 2 differ crucially in whether or not *you* can refer to the participant him- or herself. In experiment 1, participants interpreted the reports as unaddressed listeners who are excluded as possible referents of pronouns. Following Wechsler's theory, this implies that the meaning of all pronouns, including *you*, must be determined via an inferential process about another person's self-ascription. From this we can derive the prediction that there should not be systematic differences in accuracy between *you* and the other pronouns in experiment 1. This is confirmed by our empirical results that show a relatively uniform behavior of all pronouns in direct and indirect speech in an observer setting.

By contrast, in experiment 2, participants are addressees of the reports. Consequently, *you* in indirect speech refers to the participant him- or herself and is to be interpreted via self-ascription. Since self-reference is supposed to be easier than other-reference, the prediction is that children make less mistakes in interpreting *you* than *I* or *he/she* in indirect speech. Indeed, five- and nine-year-old children showed the highest accuracy rate for indirect *you*. In direct speech, second-person *you* refers not to the current addressee, the participant, but to the addressee of the reported speech context. Therefore it should be interpreted via an inferential process about someone else's self-ascription parallel to the other pronouns.

However, it turns out that the accuracy of direct *you* is not, as expected, comparable to that of direct *I* and direct *he/she*, but significantly lower. A possible explanation for this finding is that children in their role as addressees have the tendency to self-ascribe *you* even if the pronoun is embedded in a direct speech report. This explanation resembles De Hoop and Tarenskeen's (2014) suggestion on how people interpret generic *you* in a sentence such as (8).

(8) Well, what can you do about it?

De Hoop and Tarenskeen propose a two-stage interpretation process: First, people interpret the pronoun *you* as isolated lexical item via self-ascription. Second, they take the linguistic context into account, which triggers a shift to a generic interpretation of *you*. Transferred to the interpretation of *you* in direct speech, this would mean that children self-ascribe *you* in direct speech, but – in contrast to adults – are unable to revise their initial incorrect interpretation in view of the linguistic cues that signal direct speech (Epley et al., 2004; Trueswell, Sekerina, Hill, & Logrip, 1999). The fact that even nine-year-old children mostly fail to overrule the self-ascription of *you* suggests that it is a highly automatic mechanism or a strong bias. However, *you* does not exert a self-ascriptive force in all circumstances. When reference to the self is not an option, like in experiment 1, self-ascription of the second-person pronoun does not seem to be the default.

The above considerations demonstrate that the unusual behavior of *you* in experiment 2 can be explained by Wechsler's *de se* theory of pronouns. In turn, our findings thus provide empirical support for Wechsler's theory. The results indicate that children up until the age of nine use a *de se* strategy for interpreting *you* as addressees, which they also mistakenly apply when being confronted with direct speech reports.

While there is a growing body of research into adults' processing of second-person pronouns (Andeweg et al., 2013; Borghi, Glenberg, & Kaschak, 2004; Brunyé et al., 2009, 2011; Ditman et al., 2010; Sato & Bergen, 2013), comparable research on child language processing is still missing. Considering the results of the current study, further investigations about children's processing of sentences and texts with *you* as compared to *I* and *he/she* seem very promising.

5.4.3 Production

In the production part of experiment 2, participants produced speech reports in an interactive setting. They received information from one puppet and had to transmit it to another puppet. Crucially, participants were instructed to make their utterance as easy as possible for their addressee. From experiment 1 and the comprehension part of experiment 2, we know that listeners find it easier to interpret pronouns that pick out referents via their current speech-act roles in the actual reporting context. Assuming that adults take the perspective of the addressee into account, we predicted that they would produce mainly addressee-friendly indirect speech reports. This is confirmed by our results. With a few exceptions, adults produced either framed indirect speech (e.g., *Mimi said that I get the crocodile*) or indirect descriptions (*I get the crocodile*). Their preference for framed or unframed indirect reports depended on the test order. Adults who were exposed to framed reports in the comprehension part prior to production were primed to use more constructions with a reporting clause themselves.

We predicted that younger children produce a lot of direct speech reports because this type of report does not require deictic alterations compared to the original utterance.

With increasing age, we expected to find more indirect speech reports, which are addressee-friendly but more difficult to produce for the reporting speaker. The production results of the children only partly confirm these predictions.

Our analysis shows that 37% of children's reports are direct speech reports. The use of direct speech in this discourse context is not problematic as such, even though it increases the processing demands for the addressee. However, the majority of children's direct speech reports did not contain a reporting frame (*Mimi said*) or a paralinguistic quotation marking (like a change of pitch). This is consistent with previous findings that children often fail to signal that they are directly quoting someone else's speech (Goodell & Sachs, 1992; Hickmann, 1993; Nordqvist, 2001a; Özyürek, 1996). A surprising finding is that children as old as nine still produce unframed direct speech reports. An explanation for this finding is that these children leave out the reporting clause assuming that the information, who the reported speaker is, is sufficiently clear in the context. Thereby they disregard that an unframed direct speech report is highly ambiguous in this situation. If an addressee hears a sentence like "You get the crocodile" with no quotation signal, he is likely to misinterpret the pronoun *you* as referring to himself instead of the addressee of the original context of utterance.

Approximately 60% of children's reports are indirect speech reports. The use of indirect speech requires children to re-orient pronouns from the original speaker's perspective to their own current perspective as reporting speaker. This involves deictic transformations such as changing a second-person pronoun in the original utterance to a first-person pronoun in the indirect report. The high number of indirect speech reports in our data shows that at the age of five children are already able to perform these deictic adjustments. This is in line with findings from Tanz (1980). Crucially, our prediction that older children produce more reports that are deictically oriented towards the reporting context was not confirmed by the data. In fact, there was no significant difference in the number of addressee-friendly reports produced by five- and nine-year-old children. This result is surprising considering that we found a continuous improvement with age in children's comprehension of speech reports.

In sum, the production data of the children is puzzling in two respects: (i) five- and nine-year-old children produced a high number of ambiguous unframed direct reports, and (ii) older children were not better than younger children in taking the addressee's perspective into account when designing an utterance. These findings cast doubt on whether all children understood the task correctly. It is possible that some children interpreted the game as a sort of 'Chinese whispers' or 'Telephone', where the message needs to be passed on as faithfully as possible. These children might have opted for a verbatim rendition of the original message even though they would have been capable of producing indirect speech reports. If this is the case, the production results do not so much reflect children's skills to take the addressee's perspective into account, but rather their understanding of the task.

Only 0.6% (n = 9) of all speech reports produced in the production part were direct-indirect mixes. In most cases, participants used the syntax of indirect speech, but did not adjust the pronouns to the reporting context. This confirms our observation from the corpus study

in chapter 3 that the mixing of direct and indirect speech is relatively rare in production. In comparison, direct–indirect mixes occur far more frequently in comprehension. In the comprehension part of experiment 2, participants systematically misinterpreted direct as indirect speech and vice versa in 25.9% of the cases, with a decrease of mixing mistakes with age (five-year-olds: 37.6%, nine-year-olds: 24.5%, adults: 5.4%). Our findings suggest that children are able to produce correct direct and indirect speech reports, but often fail to interpret these forms correctly when used by someone else. This production–comprehension asymmetry in the acquisition of the direct–indirect distinction ties in with the observation that in child language acquisition, correct production sometimes precedes correct comprehension (Chien & Wexler, 1990; Hendriks & Spender, 2006; Hendriks et al., 2007; Hendriks, 2014).

5.4.4 Relation between comprehension and production

We tested the relation between comprehension and production in two respects.

First, we were interested in whether children’s ability to interpret pronouns in speech reports is related to their production of speech reports. We found that children who produced more addressee-friendly indirect speech reports, were significantly better in pronoun interpretation in the comprehension part. This suggests that similar skills might be involved in the production and comprehension of speech reports. In particular, we speculate that perspective-taking might be the connecting factor. When producing addressee-friendly reports, children need to take the perspective of the addressee into account. When interpreting pronouns in direct speech, children need to shift from the reporting to the reported speaker’s perspective. However, note that in the model explaining children’s production of addressee-friendly reports, the accuracy score in the comprehension part was not a predictor. This could be related to the methodological problem pointed out above that some children might have misunderstood the production task.

Second, we tested whether the order of the comprehension and production part influences children’s performance. The underlying assumption here is that a preceding comprehension or production part serves as a mini-perspective-taking training: It allows children to experience the interaction from a different standpoint and raises their awareness for the different perspectives involved in the interaction (e.g., the perspective of the addressee or reported speaker). We formulated two hypotheses on how the test order influences participants’ performance. The first hypothesis is that children tested with a preceding comprehension part produce more addressee-friendly reports that are oriented towards the reporting context. The second hypothesis is that children who are tested with a preceding production part make fewer mistakes in the comprehension part.

The results disconfirm both hypotheses: the test order (production first, comprehension first) neither predicted children’s production of addressee-friendly reports nor their accuracy of pronoun interpretation. A possible explanation for this lack of improvement is that the different perspectives were not explicitly pointed out. A training study by Lohmann

and Tomasello (2003) indicates that simply experiencing divergent perspectives is not sufficient for bringing about a training effect. In their study, children were exposed to deceptive objects (e.g., a pen that looks like a flower) and received either verbal or non-verbal explanations about their deceptive nature. Lohmann and Tomasello's results suggest that language is a necessary ingredient to structure children's perspectival experience and to improve their false belief understanding. In the context of our study, an explicit perspective-taking training could for instance consist in illustrating by concrete examples that the meaning of deictic pronouns crucially depends on the communicative context with its particular distribution of communicative roles (speaker, addressee, other).

5.5 CONCLUSION

Experiment 2 has provided new insights into children's and adults' comprehension and production of referring expressions in speech reports. By making the experimental setting more realistic and placing the participant inside of the interaction, children's accuracy of interpreting first- and third-person pronouns in direct speech improved significantly compared to experiment 1.

Experiment 2 has also revealed that processing second-person pronouns from the perspective of the addressee is special. Following Wechsler (2010), we suggested that children interpret *you* via a mechanism of self-ascription. Interestingly, children also thought that *you* referred to them when the pronoun was embedded in a direct quotation and therefore referred to the original addressee instead. This suggests that self-ascription of the second-person pronoun is the default interpretation. If the linguistic context indicates a different interpretation, like in the case of quoted *you* or generic uses of *you*, costly revisions of the initial *de se* interpretation are required.

In the production part of experiment 2, adults produced mainly indirect speech reports, which are anchored in the actual reporting context. This is also the preferred context of evaluation for listeners. Direct speech reports, which require a shift to the reported context, are dispreferred in this kind of communicative situation. To describe it with Bühler's (1934) terms, both speaker and listener prefer *demonstratio ad oculos*, i.e., pointing to things in the shared perceptual space of the here and now, to *Deixis am Phantasma* 'deixis in the imagination', i.e., reference to absent entities or – applied to the issue at hand – reference to people via their past speech-act roles. To put it pointedly, why refer to a present person with a quoted *he* or *she* rather than a *you*?

Experiments 1 and 2 have shown that it is easier for children and adults to process pronouns embedded in indirect speech reports. However, we should not jump to the conclusion that the interpretation of pronouns is always cognitively less demanding in indirect speech as opposed to direct speech. In both experiments, we tested participants' comprehension of reports in a specific kind of discourse context, which is focused on the exchange of information. In other communicative situations such as narratives, speakers prefer to use direct

speech, which allows for a more vivid, engaging presentation of speech. Listeners might also find it easier to interpret direct as compared to indirect speech reports in a narrative context. In experiment 3, the speech reports will be presented as part of a fictional narrative. This way we can test whether the type of discourse context (information exchange vs. narrative) has an impact on children's and adults' processing of reported speech.

EXPERIMENT 3: TESTING THE COMPREHENSION OF REPORTED SPEECH IN A NARRATIVE CONTEXT

6.1 INTRODUCTION AND HYPOTHESES

Experiments 1 and 2 have provided evidence that the interpretation of pronouns is cognitively more demanding in direct as compared to indirect speech. In light of these findings it is surprising that story books for beginning readers contain mainly direct speech reports (Baker & Freebody, 1989). Two features of direct speech could explain this preference: (i) the greater vividness of direct speech and (ii) the presentation of speech from the point of view of the character.

First, direct speech is more vivid and lively than indirect speech (Tannen, 2007; Wierzbicka, 1974). In contrast to indirect speech, direct speech allows the reporting speaker to demonstrate the emotional state, accent and voice of a person, for instance by the use of prosody or expressive elements (Banfield, 1973; Clark & Gerrig, 1990).

Readers view the characters speaking directly, as in play, and thus presumably experience a greater sense of immediacy or proximity to the characters. (Bortolussi & Dixon, 2003, p. 202)

The vividness of direct speech could be particularly important in connection with children because it might attract their attention and facilitate “transportation” (Gerrig, 1993) to the fictional world.

Second, stories – especially those for children – typically focus on the activities of the characters and have a minimally intrusive and almost invisible narrator (Baker & Freebody, 1989; Banfield, 1982). Direct speech has the advantage that it allows the narrator to step back and create the illusion that the characters are speaking for themselves. Because the focus of attention in a narrative is on the characters, their perspective is very salient to listeners and readers. Therefore we expect that a shift to the character’s perspective, as required in direct speech, should be easier in the context of a narrative than in the information exchange contexts of experiments 1 and 2.

In the design of experiment 3, we tried to simulate children's natural picture-book experience. In Dutch families, picture-book reading is a common practice. A survey indicates that 60% of Dutch parents read daily to their children, usually between 5 and 15 minutes a day (Duursma, 2014). On top of that, picture storybooks are read aloud almost every day in Dutch primary schools (Pool, 2012). Typically, an adult is reading from a book while the child is simultaneously listening to the story and looking at the pictures. In addition, book-reading is often associated with guiding behavior on the part of the adult who, for instance, asks the child questions about the story or draws the child's attention to certain aspects of a picture (Duursma, 2014; Vandermaas-Peeler, Nelson, Bumpass, & Sassine, 2009; Vandermaas-Peeler, Sassine, Price, & Brillhart, 2011). We incorporated both aspects into the design of experiment 3. Participants listen to a (prerecorded) narrative illustrated by pictures and occasionally have to answer comprehension questions about it. The responses to these questions provide insights into participants' interpretation of pronouns in direct and indirect speech.

Since no already existing picture book could satisfy our specific experimental requirements (e.g., concerning the number and type of speech reports, and the accompanying pictures), we decided to construct one ourselves. Based on Baker and Freebody's (1989) study on children's books, our story features two main protagonists, whose non-verbal and, crucially, verbal actions are reported by a third-person backgrounded narrator. As typical for children's stories, the utterances of the characters are concerned with observable features of the here-and-now of the scenes, namely different objects.

We have two hypotheses. The first hypothesis is that pronoun interpretation is easier in direct than in indirect speech in the narrative context of experiment 3. This is based on the finding that narratives favor the use of direct speech reports (Baker & Freebody, 1989; Vincent & Perrin, 1999). Not only do children's books contain mostly direct speech, but children themselves also produce predominantly direct speech in narrative contexts (see 3.2.4). Similar to production, we expect that children also prefer direct speech reports in comprehension. In our experiment, this would be evident from a higher accuracy of pronoun interpretation in direct speech as opposed to indirect speech. Comparing the results from experiment 1 and experiment 3, we expect children to make fewer mistakes for direct speech items in the narrative context of experiment 3 as compared to the information exchange context of experiment 1.

The second hypothesis is that the participants remember information better when it is presented in direct rather than in indirect speech. This hypothesis is inspired by studies that suggest that direct speech leads to a more vivid, perceptually engaging mental representation of a situation (Stites et al., 2013; Yao et al., 2011, 2012; Yao & Scheepers, 2011). This could enhance the memory for both content and form of the reported utterance. Groenewold, Bastiaanse, Nickels, Wieling, et al. (2014) found that Dutch adults with and without aphasia remembered the content of narratives better when they contained direct speech as compared to indirect speech. A study by Eerland et al. (2013) indicates that people memorize the exact wording better in the case of direct speech. So far, the effect of report type on children's mental representation of a situation has not been investigated.

6.2 METHOD

6.2.1 Participants

The participants of this study were 42 monolingual Dutch-speaking children between 4;1 and 7;2 years of age (see table 1). The data of one additional child was not saved due to technical problems. The participating children were recruited from an elementary school in the North of the Netherlands. Written parental consent was obtained prior to the experiment. Children received a small reward (a sticker) for participating. In addition, 20 adult native speakers of Dutch – mostly students – participated without compensation. All participants were tested individually in a quiet room at the school or university.

Table 1: Participants

Age group	Mean age	Range	Number	Gender (f/m)
4	4;7	4;1–5;0	21	8/13
6	6;11	6;8–7;2	21	13/8
adults	22	19–30	20	11/9

6.2.2 Stimuli and Procedure

The experiment has been built as an Android application and was presented to participants on a touchscreen tablet. In its design, we tried to simulate children's everyday experience of picture book reading. Participants listened to a story told by a male third-person narrator and saw illustrating pictures. The experiment consists of four parts that all form one coherent narrative: the introduction phase, the gender pre-test, the reported speech test and the memory test.¹

Introduction phase

In the beginning of the story, the narrator introduces the two main protagonists: a monkey girl called Anita Aap 'Anita Monkey' and an elephant boy called Oscar Olifant 'Oscar Elephant' (see fig. 1a). The participants were asked questions about the names and gender of the protagonists (e.g., "Who is Anita Aap?" and "Who is a boy?") and gave their answer by touching one of the highlighted protagonists (see fig. 1b). In case participants responded incorrectly, they received negative feedback from the narrator, who would subsequently ask the question again. Almost all participants answered these initial comprehension questions correctly on the first trial.

¹ The complete story and all pictures are available in Appendix C: Experiment 3. An online version of experiment 3 can be found at <http://www.philos.rug.nl/cgm/story-demo/> (Google Chrome required).

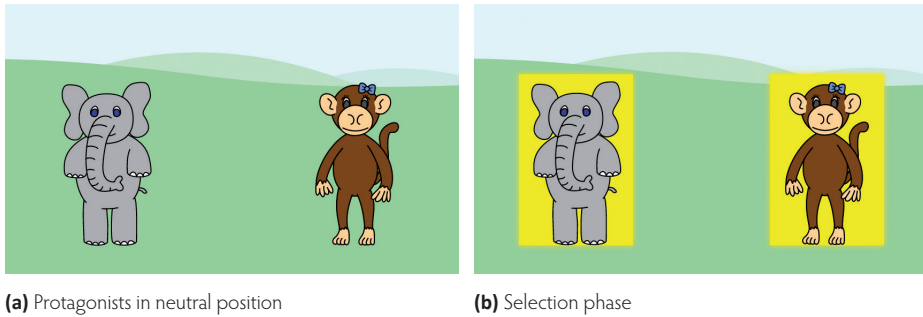


Figure 1: The protagonists Oscar Olifant and Anita Aap

Gender pre-test

After the introduction of the protagonists, we tested whether participants are able to use the gender feature of third-person pronouns as a cue for reference identification. The story continues with the narrator telling the participant that Anita and Oscar are best friends and live next to each other in two houses. One day, Oscar and Anita wake up early and, as always, start the day with a morning workout. Four sentences follow that describe who did which exercise. Two of these sentences contain the masculine pronoun *hij* 'he' (see (1)) and two the feminine pronoun *zij* 'she' (see (2)).

- (1) Hij ging touwtje springen.
'He skipped rope.'
- (2) Zij maakte een handstand.
'She did a handstand.'

The pictures accompanying these sentences show Anita and Oscar in a neutral position (see fig. 1). After each sentence, the participants had to select whether Anita or Oscar performed a particular action, for instance, doing a handstand.

Speech report test

The following part of the narrative contains 24 speech reports. The narrator describes that Oscar and Anita go for a walk after their morning workout. On their trip, Anita and Oscar come across 24 objects at different locations such as in a tree, in a pond and in a cave. All of these objects look exactly like things that they possess themselves. After the discovery of each object, the narrator reports Anita's or Oscar's utterance with either a direct or indirect speech report.

To get an impression of the test procedure, consider the example that Oscar and Anita discover an object hanging in a tree. In this case, the participants would hear the text presented in (3), accompanied by the pictures presented in figure 2.

- (3) [fig. 2a] Oscar en Anita keken omhoog en zagen ook iets in de boom hangen. Anita Aap zei tegen Oscar Olifant, “Jij hebt ook zo’n rugzak”. Wie heeft ook zo’n rugzak? [fig. 2b]
 ‘[fig. 2a] Oscar and Anita looked up and saw that there was also something hanging in the tree. Anita Aap said to Oscar Olifant, “You have a backpack like that too”. Who has a backpack like that too?’ [fig. 2b]’

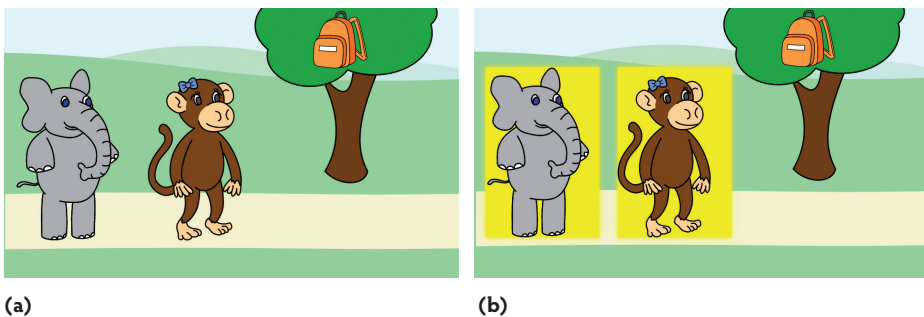


Figure 2: Examples of pictures accompanying the story

Participants had to answer the narrator’s question *Who has a backpack like that too?* by touching one of the highlighted protagonists on the screen (correct answer in (3): Oscar Olifant). The software measures the accuracy of pronoun interpretation.

The experiment contains four kinds of speech reports: direct speech and indirect speech with reference to either the reported speaker or his or her addressee (see examples (4) to (7)).

- (4) Direct speaker: Anita Aap zei tegen Oscar Olifant, “Ik heb ook zo’n voetbal”.
 ‘Anita Aap said to Oscar Olifant, “I have a football like that too”’
 (5) Direct addressee: Anita Aap zei tegen Oscar Olifant, “Jij hebt ook zo’n auto”.
 ‘Anita Aap said to Oscar Olifant, “You have a car like that too”’
 (6) Indirect speaker: Anita Aap zei tegen Oscar Olifant dat zij ook zo’n hoed heeft.
 ‘Anita Aap said to Oscar Olifant that she has a hat like that too.’
 (7) Indirect addressee: Anita Aap zei tegen Oscar Olifant dat hij ook zo’n klok heeft.
 ‘Anita Aap said to Oscar Olifant that he has a clock like that too.’

The direct speech reports contain a first- or second-person pronoun. In direct speech, the first-person pronoun *ik* ‘I’ refers to the speaker (Anita in (4)) and the second-person pronoun *jij* ‘you’ to the addressee of the reported speech context (Oscar in (5)). In indirect speech, the narrator needs to use a third-person pronoun (*hij* ‘he’ or *zij* ‘she’) to establish reference to the story protagonists. Syntactically, the third-person pronoun in the complement clause refers either to the referent that is the subject of the matrix clause, i.e., the reported speaker (Anita in (6)), or to the referent that is the object of the matrix clause, i.e., the addressee (Oscar in (7)).

In the linguistic materials, we balanced the number of direct and indirect speech reports, the reference to speaker or addressee, and whether Oscar or Anita were the subject or object of the reporting clause. In total, participants listened to 6 speech reports of each of the four types (direct speaker, direct addressee, indirect speaker, indirect addressee). In half of the reports Anita was the reported speaker, like in examples (4) to (7), in the other half Oscar was.

All participants listened to the story in the same linear order, with the exception of the speech reports (and corresponding test questions) whose order was randomized for each participant. Since all 24 speech reports are associated with a specific object, also the location at which the protagonists encounter a certain object differs between participants. For instance, a participant might see a clock instead of a backpack hanging in the tree and hear a speech report of the indirect addressee type like (7). The spatial position of Anita and Oscar (left, right) in the pictures is counterbalanced.

Experiment 3 differs in several respects from the two previous experiments. Firstly, the story includes only two instead of three characters. The reason for this is that including a third story character would make it impossible to unambiguously refer to the characters with a pronoun in indirect speech. From the point of view of the uninvolved narrator, all protagonists are third persons and therefore need to be referred to with a third-person pronoun in indirect speech. We introduced a gender difference to distinguish between the two protagonists, but an additional character could not be uniquely singled out with a masculine or feminine third-person pronoun.

Secondly, and related to the first point, we did not include all types of singular pronouns (first-person, second-person, third-person) in both direct and indirect speech, but tested only the interpretation of *I* and *you* in direct speech and *he* and *she* (referring to speaker or addressee) in indirect speech. The reason for this is that if we included first- and second-person pronouns in indirect speech, this would imply that the narrator and his addressee, the participant, would be referents of pronouns as well. While this is in principle possible, it is rather unusual that a story character refers to the narrator who is typically an invisible presenter of the events. In addition, a study by Murphy (1986) showed that children find it awkward to be directly addressed by a story character. In order to create an ecologically valid narrative context, we therefore decided to limit reference to the story characters.

Thirdly, the test sentences in experiment 3 mention the addressee in the reporting clause (e.g., *Anita Aap said to Oscar Olifant...*). We opted for this change because in indirect speech pronominal reference to the addressee would be awkward or even infelicitous if there is no linguistic antecedent (cf., *Anita Aap said that he...*).

After having discovered 23 objects at different locations, Anita and Oscar notice a strange sound. They follow the sound and discover a snoring dog leaning against a tree. Next to the dog, they find another object that looks exactly like one of theirs. A speech report, such as (4), (5), (6) or (7), follows. The story continues like this:

Suddenly Anita and Oscar realized: These WERE their things! But how was this possible? Did the dog steal their things? They woke the dog and asked him: “Did you steal our things?” Dog turned as red as fire and said: “Yes, I took your things. I’m very sorry. I just wanted to play.” Anita and Oscar thought about it for a while. “We are not angry”, said Anita. “But then you have to bring all our things back”, said Oscar. Dog said: “I promise to bring everything back immediately.”

Dog keeps his word and gathers all objects in the garden of Anita and Oscar. But now Dog has a problem because he does not remember which objects belong to Anita and which ones to Oscar. The narrator asks the participant to help Dog with the sorting of the objects.

Memory test

This is the story-internal motivation for the memory test. In the memory test, participants had to determine for each of the 24 objects whether it belongs to Oscar or Anita. Participants heard questions such as *Van wie is de rugzak?* ‘To whom does the backpack belong?’, accompanied by a picture that depicts Oscar and Anita with the object in question between them (fig.3).

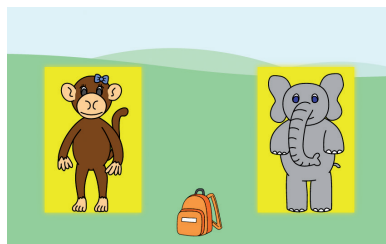


Figure 3: Example picture of memory test

We created two versions of the experiment, version (a) and version (b), and assigned participants randomly to one of them. These versions differ in the following respects: in version (a) all objects belong to the opposite story character (Anita, Oscar) than in version (b) and are associated with the opposite type of speech report (direct, indirect). The purpose of this is to control for the possibility that participants have general preferences for associating a certain object with a certain protagonist (e.g., the car with the male elephant). The complete experiment took about 10 to 15 minutes to complete.²

6.3 RESULTS

We first present the results of the gender pre-test. After that, we turn to the analysis of the speech report test, where we also compare participants' comprehension of speech reports in experiment 3 to that in experiment 1. Finally, we analyze participants' results in the memory test.

6.3.1 Gender pre-test

Figure 4 shows the results of the gender pre-test that indicates whether the experimental participants can distinguish between feminine and masculine pronouns.

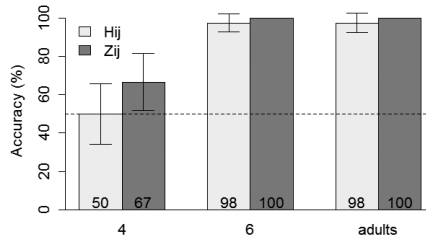


Figure 4: Percentage of correct interpretation of third-person pronouns in different age groups. Error bars indicate 95% confidence intervals.

Six-year-old children and adults were around ceiling for the interpretation of both *zij* ‘she’ and *hij* ‘he’. By contrast, four-year-old children were at chance for the masculine pronoun, $t(41)=0, p=1$, and slightly above chance for the feminine pronoun, $t(41)=2.26, p=.03$. This means that the youngest age group did not yet consistently use the gender feature to identify the referent of a pronoun.

² See Appendix C: Experiment 3 for the test sentences in version (a) and (b).

6.3.2 Speech report test: Accuracy results

Figures 5 and 6 show children's and adults' accuracy of pronoun interpretation in direct speech (fig. 5) and indirect speech (fig. 6). Recall that we distinguish between pronouns that refer to the speaker and those that refer to the addressee of the original message. In direct speech, the first-person pronoun *ik* 'I' refers to the speaker and the second-person pronoun *jij* 'you' to the addressee. In indirect speech, both *hij* 'he' and *zij* 'she' can refer to the speaking or addressed story character, depending on the subject and object of the matrix clause.

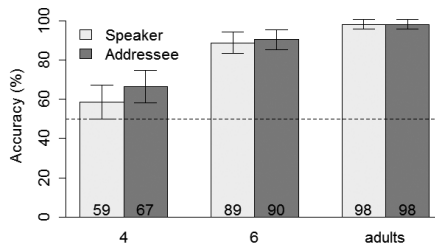


Figure 5: Percentage of correct pronoun interpretation in different age groups in direct speech. Error bars indicate 95% confidence intervals.

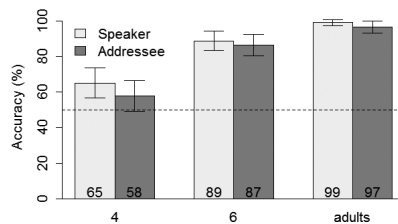


Figure 6: Percentage of correct pronoun interpretation in different age groups in indirect speech. Error bars indicate 95% confidence intervals.

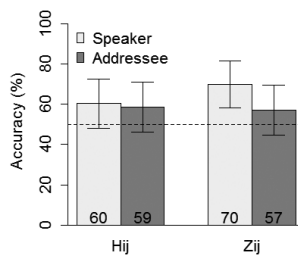


Figure 7: Zooming in on four-year-olds' accuracy of pronoun interpretation in indirect speech. Error bars indicate 95% confidence intervals.

With a mean accuracy between 87% and 90%, six-year-olds' performance was clearly above chance level for the interpretation of pronouns in both direct and indirect speech. Four-year-olds' accuracy was much closer to chance with values between 58% and 67%. Since the gender pre-test revealed significant differences between masculine and feminine pronouns in four-year-olds, we split up the indirect speech results by gender in figure 7. This reveals that only four-year-olds' interpretation of the feminine pronoun *zij* 'she' differs significantly from chance, provided that it refers to the subject of the matrix clause (speaker), $t(62) = 3.40$, $p = .001$.

We analyzed the accuracy data in the speech report test with mixed-effects logistic regression modelling. Using a procedure of model comparison, we added stepwise fixed-effect factors to the baseline model (including random intercepts for subjects). AGE and GENDER of the participants turned out to predict accuracy of pronoun interpretation. All other factors (REPORT TYPE (direct, indirect), REFERENT (speaker, addressee), PRONOUN (*I*, *you*, *he*, *she*), EXPERIENCE (1–6), SEQUENCE NUMBER (1–24), SPATIAL POSITION of the protagonists (monkey left vs. right of elephant) and VERSION (a, b)) did not improve the goodness of fit of the model. The index of concordance of the model is 0.85, which indicates that the model may have real predictive power (Baayen, 2008).

Table 2: Fixed-effects coefficients of the model fitted to accuracy of pronoun interpretation in speech reports.

	Estimate	SE	z value	p value
(Intercept)	-0.24	0.29	-0.82	.410
AGE	0.20	0.03	7.08	<.001
GENDER f vs. m	0.81	0.30	2.70	.007

Note. Model includes random intercepts for subjects.

Table 2 shows that participants' accuracy of pronoun interpretation improves with age, $p < .001$. Female participants performed significantly better than male participants, $p = .007$. A closer look at the data reveals that in the age group of the four-year-olds, girls had a mean accuracy of 67.7% ($SD = 0.47$), boys of only 58.7% ($SD = 0.49$). Among six-year-old children, girls outperformed their male peers with an accuracy of 91.3% ($SD = 0.28$) in comparison to 84.4% ($SD = 0.36$). In adults, there were no gender-related differences in accuracy.

6.3.3 Speech report test: Comparison with experiment 1

To find out whether discourse context (information exchange vs. narrative) influences the accuracy of pronoun interpretation in speech reports, we compared participants' performance in experiment 3 with the performance of same-aged participants (age groups: 4, 6 and adults) from experiment 1. In both experiments, the stimuli were presented on a tablet and participants were not potential referents of pronouns themselves. Note that the test sentences

were not identical in these two experiments³ and that the number of referential candidates differed. Participants in experiment 1 had to select a referent from a set of three potential candidates (chance level 33%), participants in experiment 3 had to choose only between two (chance level 50%). The types of reports that appeared in both experiments are direct *I*, direct *you* and indirect *he/she* (speaker). Figure 8 juxtaposes the accuracy results of these report types in experiment 1 and experiment 3.

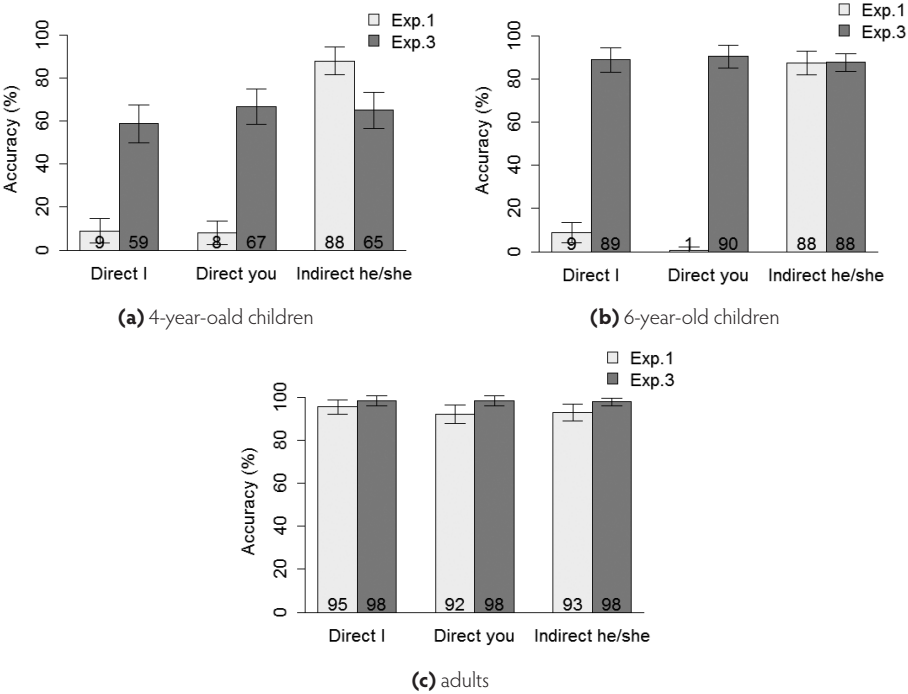


Figure 8: Percentage of correct pronoun interpretation in experiment 1 and experiment 3 in 4-year-old children (a), 6-year-old children (b), and adults (c). Error bars indicate 95% confidence intervals.

With a multiple comparisons analysis ('multcomp' package), we compared participants' accuracy in experiment 3 (narrative context) to that in experiment 1 (information exchange context). Four-year-olds were better in interpreting direct speech in the narrative context of experiment 3, both for *I* ($\beta = 3.00, z = 6.00, p < .001$) and for *you* ($\beta = 3.53, z = 6.78, p < .001$). However, four-year-olds' accuracy for indirect *he/she* was lower in experiment 3

³ The test sentences in experiment 3 mentioned the addressee in the reporting clause (e.g., *Anita Aap said to Oscar Olifant...*). Moreover, the reported proposition differed (about getting an object in experiment 1, about possessing a similar object in experiment 3).

than in experiment 1 ($\beta = -1.54, z = -3.28, p = .012$). Six-year-olds also exhibited a higher accuracy for direct *I* ($\beta = 5.39, z = 8.42, p < .001$) and direct *you* ($\beta = 8.40, z = 7.05, p < .001$) in experiment 3. Their accuracy of indirect *he/she* did not differ between experiments 1 and 3 ($\beta = -.026, z = -0.50, p = .995$). For adults, there were no significant differences between experiments 1 and 3 for all types of reports, direct *I* ($\beta = 1.08, z = 1.04, p = .896$), direct *you* ($\beta = 1.76, z = 1.75, p = .476$) and indirect *he/she* ($\beta = 1.41, z = 1.67, p = .524$).

6.3.4 Memory test

In the memory test, participants had to recall whether a certain object belongs to Anita Aap or Oscar Olifant. We determined accuracy in the memory test by comparing the memory results of each participant with their referential choice in the speech report test, which can be correct or incorrect. Participants received a score of 1 if they selected the same protagonist as in the speech report test for a particular object, otherwise they received a 0. Participants' accuracy results in the memory test are presented in figure 9.

We analyzed the memory test results with mixed-effects logistic regression. We found main effects of AGE GROUP and DISTANCE. As evident from table 3, four-year-olds had a significantly lower accuracy than adults in the memory test, $p < .01$. The accuracy of six-year-olds and adults did not differ. The factor DISTANCE indicates how many test items appeared between the mention of the object in the story and in the memory test. Since the items were randomized for each participant, also their distance score for a particular object differed. The results indicate that with increasing distance, participants' memory declined, $p < .05$. All other factors (REPORT TYPE (direct, indirect), REFERENT (speaker, addressee), PRONOUN (*I, you, she, he*), OBJECT (e.g., car, backpack) GENDER (f, m), VERSION (a, b), SEQUENCE NUMBER (1–24)) did not improve the model fit. Note that the index of concordance for this model is only 0.61, while 0.5 means random predictions and 1.0 perfect predictions (Baayen, 2008). Based on this low C value we need to be cautious concerning the reliability of the model.

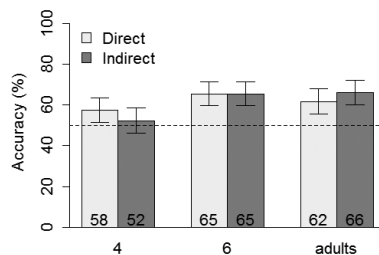


Figure 9: Percentage of correct responses in memory test (correct response means same choice of person as in speech report test). Error bars indicate 95% confidence intervals.

Table 3: Fixed-effects coefficients of the model fitted to correct responses in memory test.

	Estimate	SE	z value	p value
(Intercept)	0.91	0.17	5.36	<.001
AGE GROUP 4 vs. adults	-0.38	0.15	-2.62	.009
AGE GROUP 6 vs. adults	0.07	0.15	0.46	.649
DISTANCE	-0.01	0.01	-2.49	.013

Note. Model includes random intercepts for subjects.

6.4 DISCUSSION

In experiment 3, we tested two hypotheses. Hypothesis 1 was that in the context of a narrative, pronouns are easier to interpret in direct as compared to indirect speech. Hypothesis 2 was that children and adults remember information better when it has previously been presented in direct speech as opposed to indirect speech. The results of experiment 3 partly support hypothesis 1, but disconfirm hypothesis 2. In the following, we will discuss why our predictions were not or only partially fulfilled and consider some puzzling findings, such as gender-related differences in accuracy.

Hypothesis 1: Better accuracy in direct than indirect speech

Contrary to our predictions, children were not better in interpreting pronouns in direct speech as opposed to indirect speech in a narrative context. In fact, we found no significant differences in accuracy between direct and indirect speech. This means that reference to the characters is not easier from a character-internal viewpoint than from the point of view of the narrator. However, due to the perspectival differences between direct and indirect speech, reference to the same character is achieved with different pronouns (*I/you* vs. *he/she*). A direct comparison between direct and indirect speech in experiment 3 is therefore confounded by the deictic-anaphoric distinction, and the difference between gender-marked and non-gender-marked pronouns.

In order to find out whether children's accuracy of pronoun interpretation in direct speech improved in experiment 3, we compared the interpretation of direct *I*, direct *you* and indirect *he/she* (speaker) to that in experiment 1. Since the two experiments differ slightly (e.g., what constitutes a guessing response), care needs to be taken also in this comparison. In line with our predictions, four- and six-year-old children were significantly better in interpreting direct speech when the speech reports were part of a narrative (experiment 3) than part of an information exchange context (experiment 1). In experiment 3, the six-year-olds were almost at ceiling, with around 90% accuracy. This is remarkable when we recall that in experiment 1, eleven-year-old children had an accuracy below 50% for direct speech interpretation. This provides evidence that the pragmatic context affects the interpretation of pronouns in speech reports.

A note of caution is required for the data of the four-year-olds. In experiment 3, four-year-olds' performance in both direct and indirect speech interpretation was close to chance. This consistently poor performance could be due to higher task demands in experiment 3. Even though the stimuli in experiments 1 and 3 were presented on a tablet, the manner of presentation differs considerably. While experiment 1 consists of audio-visual animations, experiment 3 involves static pictures associated with an orally presented narrative. Animations could have two advantages over static pictures: (i) they are more engaging for children, and (ii) non-verbal cues, like mouth movement, convey information about the speech-act roles of the protagonists. In experiment 3, there were no non-verbal cues available to determine who is speaking to whom, but instead participants needed to rely solely on the linguistic information in the reporting clause (e.g., *Anita Aap said to Oscar Olifant*). This task could be cognitively demanding for four-year-old children and could be the reason why they did not perform better in identifying the referents of *I* and *you* in direct speech in experiment 3.

We also found that four-year-olds' accuracy for indirect *he/she* (speaker) was significantly lower in experiment 3 in comparison to experiment 1. In addition to the less engaging stimuli of experiment 3, this could be caused by the fact that we introduced gender differences between the animal protagonists in experiment 3. As the gender pre-test indicates, the four-year-olds were not yet able to use the gender information of third-person pronouns in their choice of referent. They were at chance for the interpretation of the masculine pronoun *hij* 'he' and slightly above chance for the interpretation of the feminine pronoun *zij* 'she'. The better performance on the feminine pronoun could be due to the fact that the gender-marking of *she* is more salient than that of *he* because *he* is also used as a gender-neutral default form in Dutch (Audring, 2009; Booij, 2002).

If we compare children's comprehension of pronominal gender in our study to that of previous studies, we find a difference in the age of acquisition. While, for instance, all English-speaking children in Arnold et al.'s (2007) study could use gender information by age four, the Dutch-speaking children in our study seem to acquire it later, between four and six. However, this does not necessarily indicate language-specific differences. The discrepancy could also be due to differences in methodology.

Gender differences

An unexpected result of this study is that girls outperformed boys in interpreting pronouns in speech reports. While it remains to be seen whether this effect will be replicated in future studies, we offer the tentative explanation that the gender difference might be related to the narrative context of experiment 3.

Support for this claim comes from both production and comprehension studies. Ely and McCabe (1993) found that English-speaking girls between the ages of four and nine produce significantly more speech reports in prompted personal narratives than their male peers (Ely & McCabe, 1993). In addition, a multitude of studies, cross-linguistic as well as

longitudinal, come to the conclusion that girls have a more positive attitude towards reading, that they read more frequently and show a better reading ability in comparison to boys (e.g., Mullis, Martin, Foy, & Drucker, 2012; Robinson & Lubienski, 2011). This gender gap in reading ability, frequency and attitude has also been documented for Dutch children (Duursma, 2014; Mullis et al., 2012). We speculate that even before children are able to read, girls enjoy listening to stories more than boys and are more proficient listeners. This could explain why they make fewer mistakes in the comprehension of speech reports in the narrative context of experiment 3.

If this turns out to be correct, more attention should be paid to the specific discourse context in which children's comprehension abilities are assessed, both in experimental studies and in educational settings. Our experiments suggest that animated interactions are equally well understood by girls and boys, whereas narrative contexts might be unfavorable to boys.

Hypothesis 2: Better memory for content presented in direct speech

Hypothesis 2 that participants have a superior memory for content presented in direct speech as opposed to indirect speech was not confirmed by the data. In fact, participants' memory for object ownership did not depend on the type of speech report in which it was mentioned. However, we found an effect of AGE GROUP and DISTANCE. Four-year-old children remembered significantly less often which protagonist owns a particular object than six-year-old children and adults. As expected, participants' memory was better the fewer items appeared between the mention of the object in the pronoun comprehension test and the memory test.

Our finding that direct speech does not improve memory for the conveyed information is not completely surprising. Even though several studies indicate that direct speech leads to a perceptually more engaging mental representation of a situation (Stites et al., 2013; Yao et al., 2011, 2012; Yao & Scheepers, 2011), it is still unclear which levels of the representation are affected by the difference in report type. Based on a series of experiments, Eerland et al. (2013) come to the conclusion that direct speech only enhances memory for the surface structure, i.e., the exact wording of the reported utterance, but not necessarily for its content. This is consistent with our findings that both children and adults did not have superior memory when a state of affairs was described in direct as compared to indirect speech.

However, it is also possible that certain aspects of our methodology are responsible for the absence of an effect. First, note that in our experiment participants' memory of person-object relations was tested. It could be the case that the mental representation of states of affairs is not affected by the contrast between direct and indirect speech, but that the representation of, for instance, actions or emotions is. Second, the information who owns a particular object was not directly relevant for the course of the story. Listeners might therefore suppress this irrelevant information during story comprehension and not create a full-fledged situation model (cf. Gernsbacher & Faust, 1991). This explanation is supported by the poor performance of the adults who answered only 64% of the memory questions correctly.

For future studies, we recommend to make the critical proposition relevant for the plot and to include different types of content (e.g., states of affairs, actions, perceptions, emotions).

6.5 CONCLUSION

Our study on children's interpretation of speech reports in narratives is part of a more general endeavor to investigate the reception of literature with psycholinguistic methods (cf. Bor-tolussi & Dixon, 2003). Since many children's books consist mainly of conversations between the characters, narrative comprehension amounts to a large extent to the understanding of speech reports (Baker & Freebody, 1989). The results of experiment 3 suggest that at the age of six – when entering school in many countries – children are able to correctly interpret both direct and indirect speech reports in a narrative. It would also be interesting to investigate at what age children understand free indirect discourse, which combines features of both direct and indirect speech (Banfield, 1973; Eckardt, 2015). Free indirect discourse is typically used in literary texts and its understanding might therefore be closely related to children's literacy development.

Experiment 3 indicates that the comprehension of speech reports in narratives is still difficult for four-year-old children. Especially their high error rate for the interpretation of *I* and *you* in direct speech is noteworthy, considering that direct speech is ubiquitous in children's books. The question arises why four-year-olds have no difficulties to interpret first- and second-person pronouns in conversations (cf. 'no report' condition of experiment 1), but struggle when characters in a story use these pronouns. At first glance, the difference between real conversations and the demonstration of conversations in a narrative seems negligible. After all, children's narratives aim at simulating real interactions by putting maximum focus on the activities of the story characters while minimizing the visibility of the narrator:

What is marked as a character's speech (e.g., 'There is a grey horse') is to be read as someone's speech and observation, as an event in some real or imagined world: what is not so marked (e.g., 'said Jack') is to be read as no one's speech or observation, and thus not an event in any real or fictional world. It is to be taken as something autonomously known about the scene and the events. (Baker & Freebody, 1989, p. 82).

If characters in a story basically speak themselves, why is it more difficult for children to interpret the utterance of a story character than the utterance of a person in a conversation? I suggest that this is related to differences in how situation models (cf. Johnson-Laird, 1983; Zwaan & Radvansky, 1998; Zwaan, 1999) are constructed in these two cases. In face-to-face conversations, an individual's mental representation of a situation relies on multiple channels of information, for instance the visual perception of objects and actions in the scene. Cues to determine who is speaking to whom are mouth-movement, bodily orientation and the

direction and quality of the sound. By contrast, when listening to a narrative, a person needs to construct a situation model purely on the basis of linguistic information. The reporting clause of a speech report (e.g., *Mary said*) usually mentions the person to whom a certain utterance is ascribed. However, the embedding of speech within speech comes with a processing cost, possibly related to a shift to a higher-order level of theory of mind.

In the light of younger children's difficulties with the understanding of narratives, it is not surprising that many children's books contain pictures. This additional visual information presumably facilitates the construction of a situation model. With respect to speech reports, comic books are an especially interesting genre. The speech and thought of comic characters is presented in bubbles that point to the designated speaker or thinker. This means that the role of speaker and addressee is indicated in the visual domain, which renders meta-linguistic frames such as *Mary thinks* or *Peter says* redundant.

For future studies, it would be interesting to compare children's comprehension of a character's speech in (a) a comic book, (b) a picture book and (c) an only linguistically presented narrative. My hypothesis is that children find it easier to understand represented conversation the more importance is placed on non-verbal visual information and the less on the linguistic domain. This predicts that comic books are easier to process than picture books and those in turn easier than books containing only text.

7

CONCLUSION

7.1 INTRODUCTION

Hardly a day goes by in which we do not hear or produce a speech report. Reporting the speech of others is such an essential part of human communication, that all natural languages we know of have a speech reporting device in their grammar. In order to become full members of a linguistic community, children therefore need to learn how to produce and interpret speech reports.

In a language like Dutch, this implies that children need to be able to distinguish direct speech from indirect speech. These two report types differ not only in syntactic respects, but more fundamentally in the perspective from which the reported utterance is presented: the reported speaker's perspective in direct speech and the reporting speaker's perspective in indirect speech. As a consequence of this perspective difference, the pronouns *I*, *you* and *he/she* can have a different meaning depending on whether they occur in a direct or indirect speech report.

Previous developmental studies indicate that children struggle to tell apart direct and indirect speech. Not only do children produce mixed reports that are not grammatical in their language (e.g., Goodell & Sachs, 1992), but they also sometimes misinterpret reports (Hollebrandse, 2007). This indicates that children have a less rigid direct–indirect distinction than adults. However, it is still unclear what linguistic or cognitive factors lie at the root of these differences between children and adults.

In this dissertation, I have addressed the following questions: Is it easier for a listener to interpret pronouns in direct or indirect speech? Do pragmatic factors influence the processing of speech reports? When do children learn to correctly interpret and produce pronouns in direct and indirect speech?

To answer these questions, I conducted three experimental studies on Dutch children's and adults' interpretation of pronouns in speech reports (chapters 4, 5 and 6). These three comprehension studies are complemented by two production studies: one a study of the spontaneous speech reports of German and Dutch children (chapter 3), the other an elicited

production study (chapter 5). The results provide converging evidence that Dutch children's acquisition of a strict direct–indirect distinction is a long process that is not completed at the age of twelve.

In the following, I will highlight the main findings of this dissertation (7.2 and 7.3), discuss the methodology (7.4) and suggest possibilities for future research (7.5).

7.2 FACTORS INFLUENCING PRONOUN INTERPRETATION IN SPEECH REPORTS

In the three psycholinguistic comprehension experiments (chapters 4, 5 and 6), I focused on the semantic difference that pronouns in direct and indirect speech are evaluated with respect to different speech contexts: the reported speech context in direct speech and the reporting speech context in indirect speech. In order to choose the correct context of evaluation, children need to be able to distinguish a direct from an indirect speech report.

The results of the experiments reveal that Dutch children's interpretation of pronouns in speech reports depends on three main factors that I will discuss in detail below: (i) the discourse context, (ii), the role in the interaction, and (iii) the presence of a linguistic antecedent.

7.2.1 Discourse context

When producing a speech report, speakers can opt for direct or indirect speech. Which form adult speakers select, depends on the type of discourse context and the function of the report. In the corpus study in chapter 3, we found evidence that Dutch and German children between the ages of 1;1 and 4;6 also adapt their choice of speech report to the communicative situation. Children preferred direct speech in contexts of pretend play and fictional narratives and indirect speech in reality-oriented contexts. Moreover, children used indirect speech more often than direct speech with an authority function, i.e., to support a statement by appealing to, for instance, their parents as authority (e.g., *Daddy said that I am allowed to do that*).

The question arises whether the type of discourse context has also an impact on children's comprehension of speech reports. In three psycholinguistic experiments, I systematically manipulated the communicative situation in which the speech reports are used. In experiments 1 and 2, the reporting speaker produced reports in order to transmit information. In experiment 3, all reports were part of a narrative, presented by a third-person narrator.

The results of experiments 1 and 2 indicate that both adults and children find indirect speech easier to process than direct speech in an information exchange context. This was evident from the fact that participants made fewer mistakes and responded more quickly when interpreting pronouns in indirect speech as opposed to direct speech. A typical error that children made was that they interpreted pronouns in direct speech like in indirect speech. This means that they did not perform a perspective shift to the reported context, but evaluated pronouns with respect to the actual reporting context. Surprisingly, even eleven-year-old chil-

dren were not yet adult-like in their interpretation of direct speech reports in an information exchange context. In the narrative context of experiment 3, however, six-year-old children were already able to correctly interpret pronouns in both direct and indirect speech.

Taken together, the results of experiments 1, 2 and 3 suggest that the discourse context influences the processing of pronouns in direct and indirect speech. But what aspect of the scenarios makes the interpretation of pronouns in direct speech more or less difficult? I argue that in different types of discourse contexts the reporting and the reported speech context exhibit different degrees of salience. In the information exchange context of experiments 1 and 2, the reporting context is more salient than the reported context. When, for instance, Dog reports to Elephant what Monkey said, Dog is clearly recognizable as speaker (by his mouth movement and voice), Elephant as addressee (by his bodily orientation towards the speaker), and Monkey as non-participant (standing at a distance and facing away from speaker and addressee). When children hear the pronouns *I*, *you* and *he* in such a situation, they link them to the most salient speaker, addressee and other person, hence Dog, Elephant and Monkey. This leads to the correct interpretation of these pronouns in non-reportative sentences and in indirect speech. Yet in direct speech, pronouns need to be evaluated with respect to the less salient reported speech context. In this situation, children seem to ignore the direct speech cues and evaluate pronouns in direct speech also with respect to the salient reporting speech context.

In the scenarios of experiments 1 and 2, listeners are therefore confronted with two conflicting factors when interpreting pronouns in direct speech: (i) the salience of the reporting speech context which ‘attracts’ speech-act participants for pronominal reference, and (ii) the syntactic (verb-second word order) and prosodic (break, change of pitch) direct speech cues that signal a context shift. While adults assign a special status to the syntactic and prosodic information, children seem to rely more on pragmatic factors such as the salience of the reporting context (cf. Elbourne, 2005; Spenader, Smits, & Hendriks, 2009). In this respect children’s interpretation of pronouns in direct speech is similar to languages like Kwaza or Danish Sign Language, in which pronouns in an otherwise shifted context can refer to the more salient actual speech-act participants. ‘Speech act participant attraction’ (Evans, 2012) thus turns out to be a more pervasive phenomenon than previously assumed and can also be successfully applied to explain (Dutch) children’s interpretation of pronouns in direct speech.

In the context of a narrative with a backgrounded third-person narrator, the reported context, i.e., the world of the story characters, is more salient than the reporting context, i.e., the narrator telling the story. When listeners hear a direct speech report in a narrative, there is therefore no conflict of cues: Both contextual salience and direct speech cues point to the reported context as context of evaluation for pronouns. This explains why children at the age of six can already interpret pronouns in direct speech correctly when the reports are part of a narrative.

What about the interpretation of indirect speech in a narrative context? We found that six-year-old children could also correctly identify the referents of pronouns in indirect speech. However, since we restricted reference to the story protagonists, we only tested children's interpretation of *he* and *she* in indirect speech. The referents of these third-person pronouns can be identified on the basis of their gender. I speculate that if we also included indirect *I* and indirect *you* in a narrative context – which we did not for reasons of ecological validity – children might incorrectly evaluate these pronouns with respect to the salient reported speech context. So, for instance when listening to the indirect report *Anita Aap said to Oscar Olifant that I have a backpack like that too*, children might select the reported speaker Anita as referent of *I* rather than the reporting speaker, the narrator.

In sum, children's interpretation of pronouns in speech reports is influenced by pragmatic aspects of the discourse context. In contrast to adults, children rely less on syntactic, lexical and prosodic cues that clearly identify a report in Dutch as direct or indirect speech. Instead, they assign a higher status to the situational embedding of the report, such as the salience of reporting and reported speech context. This makes children prone to misinterpreting pronouns in speech reports, when the contextual and linguistic cues provide conflicting information.

7.2.2 Role in the interaction

The second factor that has an impact on children's interpretation of pronouns in speech reports is the standpoint from which they process the report. In experiment 1, children listened to the report as eavesdroppers; in experiment 2, children were part of the interaction and assumed the role of the addressee in the reporting context. Our results suggest that children find it easier to interpret pronouns in direct speech from the perspective of the addressee. The reason for this could be that as a fellow interaction partner, the perspective of the reported speaker increases in salience, which facilitates the perspective shift in direct speech.

A surprising finding was that children's interpretation of the second-person pronoun *you* showed some special features when children were the addressee. In indirect speech, children's accuracy of pronoun interpretation for *you* exceeded that of *I* and *he*. In direct speech, however, children made more mistakes for *you* than for *I* and *he*. I explained this finding on the basis of Wechsler's (2010) *de se* theory of pronouns that suggests that listeners process *you* differently when they are directly addressed than when they are eavesdropping. As addressees, listeners can interpret *you* via a process of self-ascription, but as eavesdroppers, they need to determine the referent of *you* indirectly by making inferences about the self-ascriber's interpretation.

The results of experiments 1 and 2 show that the youngest children that we tested (four- to five-year-olds) were able to interpret *you* correctly both from the perspective of an addressed listener (experiment 2) and from that of an eavesdropper (experiment 1). However, in one condition, children had severe problems: when they were the addressees of the report and *you* appeared inside of a direct quotation, hence referred to the original addressee

(i.e., the current reporting speaker). In this situation, five- and nine-year-old children had the tendency to incorrectly self-ascribe *you* – and significantly more often so than difficulties to perform a perspective shift in direct speech would predict. This is the first empirical evidence that older children interpret *you* from the perspective of the addressee via a process of self-ascription, which appears to be highly automatic. Even if linguistic cues signal direct speech, children seem to find it difficult to revise their initial self-ascription.

7.2.3 Anaphora and linguistic mention

The third factor that influences children's and adults' interpretation of first-, second- and third-person pronouns in direct and indirect speech is linguistic mention. We found that it was significantly easier for both adults and children to interpret the third-person pronoun *he* in an indirect speech report like (2) than in an unembedded statement like (1).

- (1) He gets the football.
- (2) Elephant said that he gets the football.

The crucial difference between these two uses of *he* is that in (1) *he* is used deictically, i.e., to refer to an individual that is salient in the extra-linguistic context, and in (2) *he* is used anaphorically, i.e., to refer back to the subject of the matrix clause to pick up reference. In experiment 1, children and adults found it difficult to identify the referent of a deictic *he* in a situation with three male referential candidates and no additional cues such as pointing or eye gaze. But even the youngest participants of experiment 1, the four-year-olds, were able to interpret anaphoric *he* in indirect speech correctly. The reason for the ease of anaphoric *he* is that its referent is explicitly mentioned in the reporting clause, which makes him more salient.

I argue that a similar facilitating effect of linguistic mention can be observed in direct speech, even though quotations are traditionally assumed to be inert or opaque environments. In experiment 1, we found that adults and children made fewer mistakes when interpreting *I* (see (3)) as opposed to *you* (see (4)) in direct speech. I suggest that this is due to the fact that only the speaker but not the addressee is explicitly mentioned in the reporting clauses of (3) and (4).

- (3) Elephant said, “I get the football”.
- (4) Elephant said, “You get the football”.
- (5) Anita Aap said to Oscar Olifant, “You have a backpack like that too”.

Once we also included the addressee in the reporting clause, like in (5) – as we did in experiment 3 – the difference in accuracy between *I* and *you* disappeared.

Because of their similarities with anaphoric *he*, *I* and *you* in direct reports like (3) and (5) have been called “instances of anaphora, albeit indirect anaphora” (Halliday & Hasan,

1976, p. 49). Even though these deictic pronouns are not syntactically bound by the subject or object of the matrix clause, the presence of co-referential terms in the reporting clause makes it easier for listeners to process *I* and *you* in a quotation. The reason for this is that a linguistic mention makes individuals more salient and therefore more easily accessible for pronominal reference.

Compared to the processing of anaphoric pronouns, the processing of deictic pronouns has received surprisingly little attention. While the influence of linguistic factors such as topic-hood, grammatical role, order and frequency of mention is well-studied for the processing of anaphoric pronouns, similar psycholinguistic investigations on the processing of deictic pronouns in direct quotations are required. Our studies provide a first indication that linguistic factors influence not only children's and adults' interpretation of anaphoric pronouns but also their processing of deictic pronouns embedded in direct speech reports.

7.3 A PRODUCTION–COMPREHENSION ASYMMETRY?

In the analysis of Dutch and German children's spontaneous speech reports in chapter 3, we found that typically developing children are able to produce direct and indirect speech by age three. The order in which children acquire the different direct and indirect speech constructions in Dutch and German is related to syntactic complexity. While direct quotations of sounds and words are children's first speech reports, direct speech reports of complete clauses, and verb-second and verb-final indirect speech emerge several months later. To our surprise, children rarely produced speech reports that mix features of direct and indirect speech – both in natural interactions and in the production part of experiment 2.

These production results contrast sharply with our findings from the comprehension experiments. In experiment 1, for instance, even eleven-year-old children interpreted pronouns in direct speech in more than 50% of the cases like in indirect speech or non-reportative sentences. Does this mean that correct production precedes correct comprehension of direct and indirect speech? Production–comprehension asymmetries have been observed in many areas of language acquisition, for instance third-person deictic pronouns in French and anaphoric object pronouns in Dutch and English (see Hendriks, 2014, for an overview). Our results suggest that children's acquisition of the direct–indirect distinction is another case in point.

However, even though our results indicate that children's comprehension lags considerably behind their production of speech reports, the effects on everyday conversations need not be dramatic. Children's interaction partners are usually cooperative and select the form of speech report that requires the least processing effort for the addressee. This is confirmed by the results from the production part of experiment 2: Adults produced mainly indirect speech reports in an information exchange context, the type of report that even four-year-olds can process without difficulty in this situation. So there is reason to believe that children are equally good in producing and interpreting pronouns in speech reports if a pragmatically suitable discourse context is provided.

Despite these worries about ecological validity, the experimental studies revealed differences between children and adults that probably go largely unnoticed in everyday life. Adults' direct–indirect distinction appears to be more rigid than that of children because it is based on the clear syntactic, lexical and prosodic differences that distinguish direct and indirect speech in Dutch. Children's greater reliance on contextual factors makes them more prone to misunderstanding direct as indirect speech or vice versa.

7.4 METHODOLOGICAL REMARKS

In all three comprehension experiments in this dissertation, participants had to identify the referents of pronouns in direct and indirect speech. The methodology differed, however, in several respects. First, in experiments 1 and 3 we presented the stimuli on a tablet, and in experiment 2 in a more realistic scenario with hand puppets. Second, when we used a tablet, the reports were accompanied by animations (experiment 1) or static pictures (experiment 3). In the following, I will discuss the advantages and disadvantages of the different methods.

Tablet vs. hand puppets

Presenting the stimuli on a tablet as opposed to with hand puppets has several advantages. The first advantage is that the experimental app automatically measures and stores values such as accuracy and decision times. Second, the touch screen offers an easy and intuitive way to select the referent of a pronoun. Even four-year-olds without prior tablet experience showed no problems with this selection procedure. Third, children enjoyed playing our experimental 'game' on a tablet and were eager to participate. It is fair to say that they were more excited about the tablet than about the hand puppets.

With the use of hand puppets and other props, we tried to simulate a more natural communicative environment, in which the participants were involved in the interaction. Even though children can in principle also be addressed in a tablet game, this kind of address is difficult for children because they are not clearly part of the same world as the speaker (cf. Murphy, 1986). We decided against using prerecorded stimuli in the puppet experiment to make the interaction between the puppets and the child more natural. However, there is a tradeoff between the naturalness and the comparability and reproducibility of the direct and indirect speech stimuli. A disadvantage of the use of puppets and other props is that some of the youngest children seemed to be intimidated by the complex setup including three experimenters, two hand puppets as well as objects and boxes. This led to a higher dropout rate than in the two tablet studies.

In sum, there are advantages and disadvantages of presenting the stimuli on a tablet or in a more realistic scenario with hand puppets. They need to be carefully considered and weighed against each other when designing future experiments.

Animations vs. static pictures

When testing the comprehension of speech reports on a tablet, we used either animations (experiment 1) or static pictures (experiment 3) to illustrate the interactions. Animations have the advantage that they provide additional non-verbal information. Children can, for instance, identify the speaker by his mouth movement. In the case of static pictures, typical for book-reading contexts, children need to rely mainly on verbal information to find out who the speaker is. Our results indicate that this is challenging for children younger than six. In order to support children's comprehension of speech reports in picture books, the reading adult can provide non-verbal information herself, e.g., by pointing to the character whose speech is presented. For experimental studies, I recommend that researchers should consider using animations more often. They are not only easier for children to understand, but also especially engaging.

7.5 FUTURE RESEARCH

In this dissertation, I have examined one specific aspect that distinguishes direct and indirect speech: the interpretation of deictic pronouns. I tested typically developing children and adults in Dutch, a language with a clearly marked direct–indirect distinction.

In future studies, it would be interesting (i) to test pronoun comprehension in speech reports in other populations and (ii) to focus on different aspects of the direct–indirect distinction. In the following, I suggest promising possibilities for further research.

Individuals with Autism Spectrum Disorder

Previous studies have shown that individuals with Autism Spectrum Disorder (ASD) have difficulties with the interpretation and production of pronouns. Children with ASD tend to reverse the pronouns *I* and *you* initially (Kim et al., 2014; Tager-Flusberg, 1994), and even adults with ASD struggle with the deictic shifting involved in pronoun use (Mizuno et al., 2011). It is still unclear what implications this has for the interpretation of pronouns in indirect speech and especially in direct speech. In direct speech, a cognitively demanding shift from the reporting speaker's perspective to the reported speaker's perspective is required, which could be especially challenging for individuals with ASD (cf. Baron-Cohen et al., 1985). In future studies, we could test children and adults with ASD on the three comprehension experiments described in this dissertation. Since in experiment 2, but not in experiments 1 and 3, the experimental participants are possible referents of pronouns, we can contrast the interpretation of pronouns that refer to the self with the interpretation of pronouns that refer to other people. This is particularly interesting because previous studies indicate that people with ASD differ from typical adults in terms of self-reference (cf. Frith & De Vignemont, 2005; Mizuno et al., 2011).

Elderly people

While I focused in this book on children and younger adults, speakers at the other end of the lifespan face different challenges. Previous studies indicate that with increasing age, adults' processing of syntactically complex sentences declines along with certain cognitive skills (Burke & Shafto, 2008). Since speech reports are frequent in everyday conversations, it is important to find out whether elderly people struggle to interpret pronouns in direct and indirect speech. In a preliminary study, Visker (2015) and Ellen¹ tested a group of elderly Dutch people with my tablet experiments 1 and 3. The results indicate that older adults tend to make more mistakes than younger adults in the cognitively demanding direct speech condition in experiment 1. Further studies are necessary to find out whether this age-related decline in performance is due to, for instance, sensory/perceptual deficits, generally slower processing, inhibition or working memory deficits.

Bilinguals

In the Netherlands, there are a couple of hundred thousand people who are bilingual Frisian-Dutch speakers. Interestingly, in the type of Frisian spoken in the Netherlands, embedded main clause constructions such as *Aap sei, hy krijt de auto* 'Monkey said, he gets the car' are grammatical in contrast to standard Dutch (De Haan, 2010; Zwart, 1997). This means that the direct-indirect distinction is less clearly syntactically marked in Frisian than in Dutch because both direct and indirect speech reports can have verb-second word order. Based on this difference, I hypothesize that when listening to Dutch speech reports, Frisian-Dutch bilinguals are more inclined to interpret direct speech as indirect speech than monolingual Dutch speakers.²

In a preliminary study, Van der Meer (2013)³ compared the pronoun interpretation of Frisian-Dutch bilingual adults and Dutch monolingual adults, using experiment 1. The preliminary results suggest that Frisian-Dutch bilinguals make significantly more mistakes than Dutch monolinguals in their interpretation of pronouns in both direct and indirect speech. This could be due to a language transfer from Frisian to Dutch. To substantiate this claim, further studies are necessary.

¹ Visker and Ellen wrote or are in the process of writing their Bachelor theses on elderly people's interpretation of pronouns in speech reports under the supervision of Jennifer Spenader (University of Groningen).

² I take 'monolingual Dutch speaker' here to mean that Dutch is the only native language of a person.

³ Van der Meer wrote his Bachelor thesis on Frisian-Dutch bilinguals' interpretation of pronouns in speech reports under the supervision of Jennifer Spenader (University of Groningen).

Individuals with aphasia

Earlier studies indicate that individuals with aphasia produce significantly more direct speech reports than non-brain damaged adults and find it easier to comprehend narrative texts with direct as compared to indirect speech reports (Groenewold et al., 2013; Groenewold, Bastiaanse, Nickels, Wieling, et al., 2014). Based on these findings, Groenewold (2015, p. 140) gives the following recommendation: “when talking to Dutch individuals with aphasia, we recommend communication partners to use direct rather than indirect speech constructions to report speech”.

However, in the light of our finding that the processing effort of speech reports depends on the discourse context, this statement probably needs to be qualified. In an information exchange context, for instance, all groups of participants that we tested found it easier to interpret indirect speech reports. In order to obtain further insights into the processing of speech reports in aphasia, we should test aphasic speakers’ comprehension of speech reports in different discourse contexts. The easily understandable tablet experiments 1 and 3, originally designed for testing children, are highly suitable for this purpose.

Other languages

Looking beyond Dutch, new insights can be gained by studying the comprehension of speech reports in languages like Kwaza or Danish Sign Language, in which the direct–indirect distinction is less clear-cut than in Dutch. Special attention should be paid to the pragmatic conditions under which pronominal reference in an otherwise shifted report is ‘attracted’ by the present speech act participants (Evans, 2012). Our findings indicate that Dutch children unquote pronouns in direct speech in situations in which the speech-act participants of the reporting speech context are highly salient. The question arises whether it is easier for children to reach an adult-like level of reported speech comprehension in a language in which the interpretation of pronouns is driven by contextual salience to a greater extent.

Other context-dependent expressions

If we go beyond the interpretation and production of pronouns in speech reports, other areas for future research open up. It is still unclear how adults and children interpret other types of context-dependent expressions in speech reports. Plank’s (1986) exploratory study indicates that German adults allow a lot of freedom in the interpretation of the spatial indexicals *here* and *there* and the temporal indexicals *today* and *tomorrow* in direct and indirect speech. This suggests that the direct–indirect distinction in German is less rigid in the domain of space and time deixis than in the domain of person deixis. Also the interpretation of speaker-dependent expressions such as epithets (e.g., *that bastard*) and sentence adverbs (e.g., *unfortunately*) allow a more flexible interpretation than traditionally assumed (Eckardt, 2015;

Harris & Potts, 2009). Future studies could compare whether adults and children have different preferences for interpreting spatial and temporal indexicals, epithets and speaker adverbs in direct and indirect speech. Furthermore, in view of our findings it would be interesting to explore what semantic and pragmatic factors make a listener evaluate an expression with respect to the reported or the reporting speaker's perspective.

Free indirect discourse

In this dissertation, I have studied the comprehension of direct and indirect speech, leaving aside a third type of reported speech: free indirect discourse. Free indirect discourse contains features of canonical direct and canonical indirect speech and is characterized by its mingling of the character's and the narrator's perspective. Because of this mixing of perspectives, it would be interesting to investigate children's and adults' processing of free indirect discourse with psycholinguistic methods. Given that free indirect discourse is mainly used as a literary technique, I expect that children's comprehension of free indirect discourse is delayed compared to their comprehension of direct and indirect speech reports. Furthermore, children's understanding of free indirect discourse is probably related to their literacy development.

As the long – and by no means exhaustive – list of possible future studies indicates, we still have a long way to go before we understand the cognitive processes underlying the comprehension and production of speech reports in different groups of people and different languages. In this light, the in-depth investigation of Dutch children's and adults' interpretation and production of pronouns in speech reports is but a small step in this direction.

APPENDICES

APPENDIX A: EXPERIMENT 1

An online version of experiment 1 including animations can be found at <http://tinyurl.com/o7bburc> (Google Chrome required).

'No report' condition

Table A.1: Test items of the 'no report' condition (order randomized)

Pronoun	Test sentence	Speaker/addressee
I	Ik krijg de paraplu.	Monkey/Elephant
I	Ik krijg de pen.	Monkey/Dog
I	Ik krijg de sjaal.	Dog/Monkey
I	Ik krijg de zonnebril.	Elephant/Monkey
I	Ik krijg het vliegtuig.	Elephant/Dog
You	Jij krijgt de hamer.	Monkey/Dog
You	Jij krijgt de gitaar.	Dog/Monkey
You	Jij krijgt de auto.	Dog/Elephant
You	Jij krijgt de kop.	Elephant/Monkey
You	Jij krijgt de vlag.	Elephant/Dog
He	Hij krijgt de lepel.	Monkey/Elephant
He	Hij krijgt de voetbal.	Monkey/Dog
He	Hij krijgt de schaar.	Dog/Monkey
He	Hij krijgt de tandenborstel.	Dog/Elephant
He	Hij krijgt het boek.	Elephant/Monkey

'Speech report' condition**Table A.2:** Test items of the 'speech report' condition (order randomized)

Report type	Pronoun	Test sentence	Speaker/addressee (reporting context)
Direct	I	Hond zei: "Ik krijg de hamer".	Monkey/Elephant
Direct	I	Olifant zei: "Ik krijg de zonnebril".	Monkey/Dog
Direct	I	Olifant zei: "Ik krijg de voetbal".	Dog/Monkey
Direct	I	Aap zei: "Ik krijg de sleutel".	Elephant/Dog
Direct	I	Hond zei: "Ik krijg de hoed".	Elephant/Monkey
Direct	You	Hond zei: "Jij krijgt de schaar".	Monkey/Elephant
Direct	You	Olifant zei: "Jij krijgt de pen".	Monkey/Dog
Direct	You	Aap zei: "Jij krijgt de roos".	Dog/Elephant
Direct	You	Olifant zei: "Jij krijgt de lepel".	Dog/Monkey
Direct	You	Aap zei: "Jij krijgt de vlag".	Elephant/Dog
Direct	He	Hond zei: "Hij krijgt de roos".	Monkey/Elephant
Direct	He	Aap zei: "Hij krijgt de auto".	Dog/Elephant
Direct	He	Olifant zei: "Hij krijgt de paraplu".	Dog/Monkey
Direct	He	Aap zei: "Hij krijgt de kop".	Elephant/Dog
Direct	He	Hond zei: "Hij krijgt het vliegtuig".	Elephant/Monkey
Indirect	I	Olifant zei dat ik de voetbal krijg.	Monkey/Dog
Indirect	I	Aap zei dat ik de tandenborstel krijg.	Dog/Elephant
Indirect	I	Olifant zei dat ik de hoed krijg.	Dog/Monkey
Indirect	I	Aap zei dat ik de vlag krijg.	Elephant/Dog
Indirect	I	Hond zei dat ik het boek krijg.	Elephant/Monkey
Indirect	You	Hond zei dat jij de kop krijgt.	Monkey/Elephant
Indirect	You	Olifant zei dat jij de sjaal krijgt.	Monkey/Dog
Indirect	You	Olifant zei dat jij de hamer krijgt.	Dog/Monkey
Indirect	You	Aap zei dat jij de pen krijgt.	Elephant/Dog
Indirect	You	Hond zei dat jij het boek krijgt.	Elephant/Monkey
Indirect	He	Hond zei dat hij de tandenborstel krijgt.	Monkey/Elephant
Indirect	He	Olifant zei dat hij de schaar krijgt.	Monkey/Dog
Indirect	He	Aap zei dat hij het vliegtuig krijgt.	Dog/Elephant
Indirect	He	Olifant zei dat hij de gitaar krijgt.	Dog/Monkey
Indirect	He	Aap zei dat hij de zonnebril krijgt.	Elephant/Dog

APPENDIX B: EXPERIMENT 2

Comprehension part

Table B1: Test sentences of the comprehension part

Order	Report	Pronoun	Test sentence
1	Indirect	I	Bobo/Lola zei dat ik de stoel krijg.
2	Direct	He/She	Bobo/Lola zei: "Hij/Zij krijgt het nijlpaard".
3	Indirect	You	Bobo/Lola zei dat jij de clown krijgt.
4	Direct	I	Bobo/Lola zei: "Ik krijg de boot".
5	Direct	You	Bobo/Lola zei: "Jij krijgt de haai".
6	Indirect	I	Bobo/Lola zei dat ik het kanon krijg.
7	Indirect	You	Bobo/Lola zei dat jij de bloem krijgt.
8	Indirect	He/She	Bobo/Lola zei dat hij/zij de schep krijgt.
9	Direct	He/She	Bobo/Lola zei: "Hij/Zij krijgt de hoed".
10	Indirect	He/She	Bobo/Lola zei dat hij/zij de brandweerman krijgt.
11	Direct	You	Bobo/Lola zei: "Jij krijgt de fiets".
12	Direct	I	Bobo/Lola zei: "Ik krijg de koe".
13	Direct	You	Bobo/Lola zei: "Jij krijgt de olifant".
14	Indirect	I	Bobo/Lola zei dat ik de ridder krijg.
15	Indirect	You	Bobo/Lola zei dat jij de slee krijgt.
16	Direct	You	Bobo/Lola zei: "Jij krijgt de geit".
17	Indirect	He/She	Bobo/Lola zei dat hij/zij de emmer krijgt.
18	Direct	He/She	Bobo/Lola zei "Hij/Zij krijgt de aap".
19	Indirect	I	Bobo/Lola zei dat ik de boom krijg.
20	Direct	I	Bobo/Lola zei: "Ik krijg de papegaai".
21	Direct	He/She	Bobo/Lola zei: "Hij/Zij krijgt het varken".
22	Indirect	You	Bobo/Lola zei dat jij de tafel krijgt.
23	Direct	I	Bobo/Lola zei: "Ik krijg het paard".
24	Direct	You	Bobo/Lola zei: "Jij krijgt de poes".
25	Indirect	He/She	Bobo/Lola zei dat hij/zij de beer krijgt.
26	Direct	I	Bobo/Lola zei: "Ik krijg het vuur".
27	Direct	He/She	Bobo/Lola zei: "Hij/Zij krijgt de kinderwagen".
28	Indirect	I	Bobo/Lola zei dat ik de bezem krijg.
29	Indirect	He/She	Bobo/Lola zei dat hij/zij de schildpad krijgt.
30	Indirect	You	Bobo/Lola zei dat jij de haan krijgt.

Note. The test sentences for the male participants included *Bobo* and *hij* 'he', the test sentences for the female participants included *Lola* and *zij* 'she'. A second test order was used, in which the first and second half of the test sentences is reversed (16–30, 1–15).

Production part

Table B2: Test sentences of the production part

Order	Pronoun	Test sentence
1	You	Jij krijgt het skateboard.
2	I	Ik krijg het hert.
3	He/She	Hij/Zij krijgt Sinterklaas.
4	You	Jij krijgt het bankje.
5	You	Jij krijgt de piraat.
6	He/She	Hij/Zij krijgt het bed.
7	I	Ik krijg de helm.
8	You	Jij krijgt de krokodil.
9	I	Ik krijg de neushoorn.
10	He/She	Hij/Zij krijgt de ladder.
11	I	Ik krijg de politieman.
12	You	Jij krijgt het schaap.
13	He/She	Hij/Zij krijgt de giraf.
14	I	Ik krijg de scooter.
15	He/She	Hij/Zij krijgt de hond.

Note. Test sentences for male participants included *hij* 'he', test sentences for female participants included *zij* 'she'. A second test order was used, in which the first and second half of the test sentences is reversed (8–15, 1–7).

APPENDIX C: EXPERIMENT 3

An online version of experiment 3 can be found at <http://www.philos.rug.nl/cgm/story-demo/> (Google Chrome required).

Table C1: Test sentences, version a (order randomized)

Report	Referent	Test sentence
Direct	Speaker	Anita Aap zei tegen Oscar Olifant: "Ik heb ook zo'n voetbal".
Direct	Speaker	Anita Aap zei tegen Oscar Olifant: "Ik heb ook zo'n boek".
Direct	Speaker	Anita Aap zei tegen Oscar Olifant: "Ik heb ook zo'n gitaar".
Direct	Speaker	Oscar Olifant zei tegen Anita Aap: "Ik heb ook zo'n zonnebril".
Direct	Speaker	Oscar Olifant zei tegen Anita Aap: "Ik heb ook zo'n pen".
Direct	Speaker	Oscar Olifant zei tegen Anita Aap: "Ik heb ook zo'n sjaal".
Direct	Addressee	Anita Aap zei tegen Oscar Olifant: "Jij hebt ook zo'n paraplu".
Direct	Addressee	Anita Aap zei tegen Oscar Olifant: "Jij hebt ook zo'n kopje".
Direct	Addressee	Anita Aap zei tegen Oscar Olifant: "Jij hebt ook zo'n auto".
Direct	Addressee	Oscar Olifant zei tegen Anita Aap: "Jij hebt ook zo'n vliegtuig".
Direct	Addressee	Oscar Olifant zei tegen Anita Aap: "Jij hebt ook zo'n lepel".
Direct	Addressee	Oscar Olifant zei tegen Anita Aap: "Jij hebt ook zo'n emmer".
Indirect	Speaker	Anita Aap zei tegen Oscar Olifant dat zij ook zo'n vlag heeft.
Indirect	Speaker	Anita Aap zei tegen Oscar Olifant dat zij ook zo'n hoed heeft.
Indirect	Speaker	Anita Aap zei tegen Oscar Olifant dat zij ook zo'n tandenborstel heeft.
Indirect	Speaker	Oscar Olifant zei tegen Anita Aap dat hij ook zo'n schaar heeft.
Indirect	Speaker	Oscar Olifant zei tegen Anita Aap dat hij ook zo'n sleutel heeft.
Indirect	Speaker	Oscar Olifant zei tegen Anita Aap dat hij ook zo'n stoel heeft.
Indirect	Addressee	Oscar Olifant zei tegen Anita Aap dat zij ook zo'n rugzak heeft.
Indirect	Addressee	Oscar Olifant zei tegen Anita Aap dat zij ook zo'n lamp heeft.
Indirect	Addressee	Oscar Olifant zei tegen Anita Aap dat zij ook zo'n schoen heeft.
Indirect	Addressee	Anita Aap zei tegen Oscar Olifant dat hij ook zo'n bezem heeft.
Indirect	Addressee	Anita Aap zei tegen Oscar Olifant dat hij ook zo'n klok heeft.
Indirect	Addressee	Anita Aap zei tegen Oscar Olifant dat hij ook zo'n skateboard heeft.

Note. Each test sentence is followed by a test question that includes the object mentioned in the test sentence (e.g., *Wie heeft ook zo'n voetbal?* 'Who has a football like that too?').

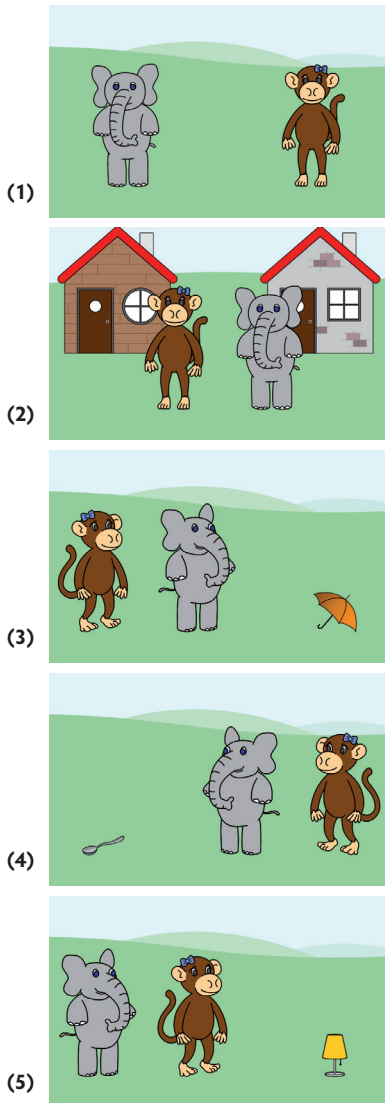
Table C2: Test sentences, version b (order randomized)

Report	Referent	Test sentence
Direct	Speaker	Oscar Olifant zei tegen Anita Aap: "Ik heb ook zo'n vlag".
Direct	Speaker	Oscar Olifant zei tegen Anita Aap: "Ik heb ook zo'n hoed".
Direct	Speaker	Oscar Olifant zei tegen Anita Aap: "Ik heb ook zo'n tandenborstel".
Direct	Speaker	Anita Aap zei tegen Oscar Olifant: "Ik heb ook zo'n schaar".
Direct	Speaker	Anita Aap zei tegen Oscar Olifant: "Ik heb ook zo'n sleutel".
Direct	Speaker	Anita Aap zei tegen Oscar Olifant: "Ik heb ook zo'n stoel".
Direct	Addressee	Anita Aap zei tegen Oscar Olifant: "Jij hebt ook zo'n rugzak".
Direct	Addressee	Anita Aap zei tegen Oscar Olifant: "Jij hebt ook zo'n lamp".
Direct	Addressee	Anita Aap zei tegen Oscar Olifant: "Jij hebt ook zo'n schoen".
Direct	Addressee	Oscar Olifant zei tegen Anita Aap: "Jij hebt ook zo'n bezem".
Direct	Addressee	Oscar Olifant zei tegen Anita Aap: "Jij hebt ook zo'n klok".
Direct	Addressee	Oscar Olifant zei tegen Anita Aap: "Jij hebt ook zo'n skateboard".
Indirect	Speaker	Oscar Olifant zei tegen Anita Aap dat hij ook zo'n voetbal heeft.
Indirect	Speaker	Oscar Olifant zei tegen Anita Aap dat hij ook zo'n boek heeft.
Indirect	Speaker	Oscar Olifant zei tegen Anita Aap dat hij ook zo'n gitaar heeft.
Indirect	Speaker	Anita Aap zei tegen Oscar Olifant dat zij ook zo'n zonnebril heeft.
Indirect	Speaker	Anita Aap zei tegen Oscar Olifant dat zij ook zo'n pen heeft.
Indirect	Speaker	Anita Aap zei tegen Oscar Olifant dat zij ook zo'n sjaal heeft.
Indirect	Addressee	Oscar Olifant zei tegen Anita Aap dat zij ook zo'n paraplu heeft.
Indirect	Addressee	Oscar Olifant zei tegen Anita Aap dat zij ook zo'n kopje heeft.
Indirect	Addressee	Oscar Olifant zei tegen Anita Aap dat zij ook zo'n auto heeft.
Indirect	Addressee	Anita Aap zei tegen Oscar Olifant dat hij ook zo'n vliegtuig heeft.
Indirect	Addressee	Anita Aap zei tegen Oscar Olifant dat hij ook zo'n lepel heeft.
Indirect	Addressee	Anita Aap zei tegen Oscar Olifant dat hij ook zo'n emmer heeft.

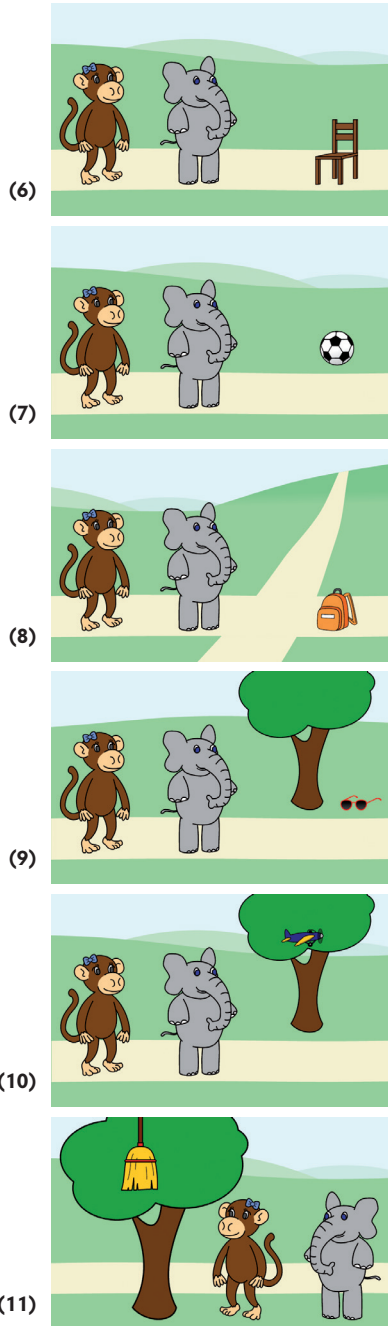
Note. Each test sentence is followed by a test question that includes the object mentioned in the test sentence (e.g., *Wie heeft ook zo'n voetbal?* 'Who has a football like that too?').

Narrative (translation below)

Explanatory note: The appearance of a certain picture is indicated in round brackets (e.g., (1)). ‘[TS+Q]’ indicates that a test sentence is presented, followed by the respective question (see Tables C1 and C2). The order of the test sentences is randomized and with it the kinds of objects that Monkey and Elephant discover at different locations. The spatial position of Monkey and Elephant (left, right) is also randomized. The following pictures are therefore only examples. For a more authentic experience of experiment 3, please have a look at the online version.

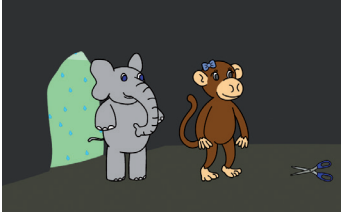


(1) Dit zijn Anita Aap en Oscar Olifant. Anita Aap is een meisje en Oscar Olifant is een jongen. Wie is Anita Aap? Wie is Oscar Olifant? Wie is een meisje? Wie is een jongen? (2) Oscar Olifant en Anita Aap zijn beste vriendjes. Ze wonen naast elkaar in twee huisjes. Oscar Olifant woont in een huisje van steen en Anita Aap woont in een huisje van hout. Ze zien elkaar elke dag en samen beleven ze leuke avonturen. (1) Op een dag werden Anita Aap en Oscar Olifant vroeg wakker. Zoals altijd begonnen ze de dag met ochtendgymnastiek. Oscar en Anita deden allebei iets: Zij/Hij maakte een handstand. Wie maakte een handstand? Zij/Hij liep een rondje om de huisjes. Wie liep een rondje om de huisjes? Zij/Hij ging touwtje springen. Wie ging touwtje springen? Zij/Hij maakte een koprol. Wie maakte een koprol? Na het sporten, besloten Oscar Olifant en Anita Aap samen een wandeling te gaan maken. Het was een mooie dag en het was 's ochtends al lekker warm. Ze liepen op blote voeten door het gras. (3) Daar zagen ze iets in het gras liggen. [TS+Q] (4) Er lag nog iets in het gras. [TS+Q] (5) Anita en Oscar liepen en paar passen verder en kwamen nog iets tegen. [TS+Q]

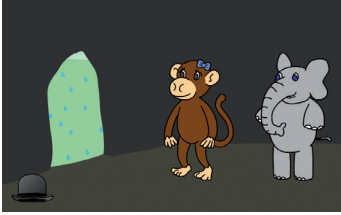


Oscar en Anita liepen verder. Ze kwamen uit op een pad en volgden het een stukje. (6) Opeens zagen ze iets op het pad liggen. [TS+Q] (7) En er lag ook nog iets naast het pad. [TS+Q] Anita en Oscar liepen nog een stukje verder. (8) Bij het kruispunt lag nog iets op de grond. [TS+Q] Ze gingen verder. Anita Aap floot een liedje en Oscar Olifant toeterde mee met zijn slurf. Ze hadden het heel gezellig met elkaar. (9) Ineens zagen ze iets onder een boom liggen. [TS+Q] (10) Oscar en Anita keken omhoog en zagen ook iets in de boom hangen. [TS+Q] (11) Ze liepen om de boom heen en zagen aan de andere kant ook nog iets in de boom hangen. [TS+Q]

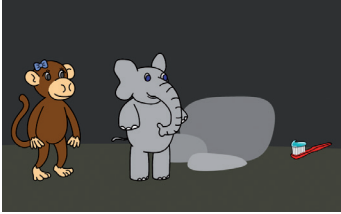
(12)



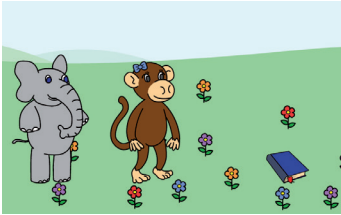
(13)



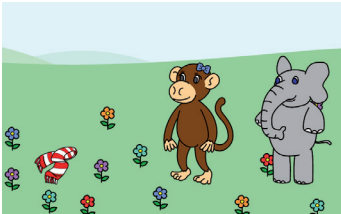
(14)



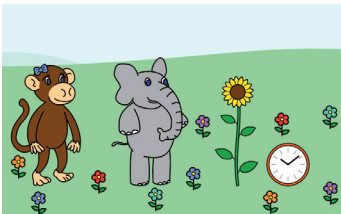
(15)



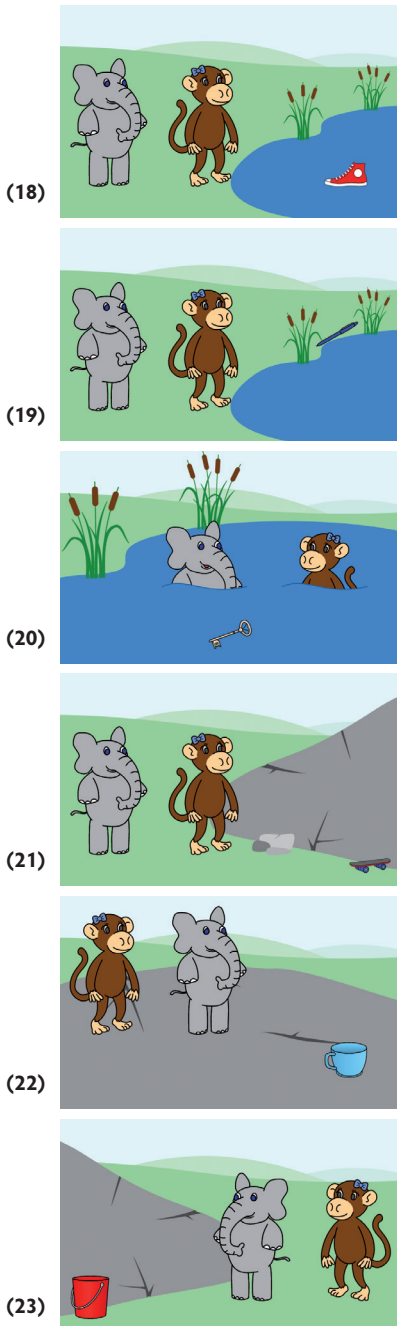
(16)



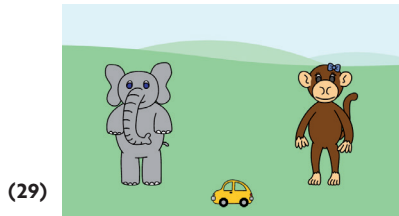
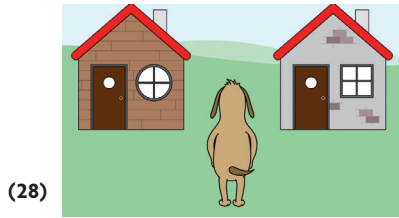
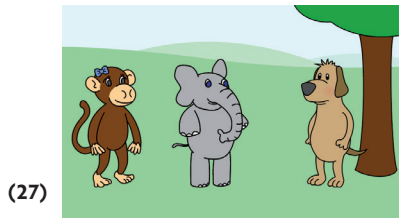
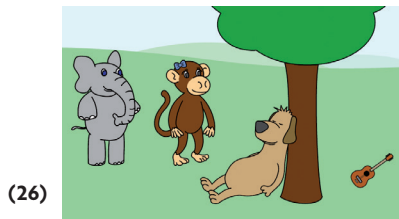
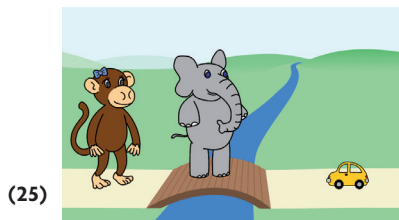
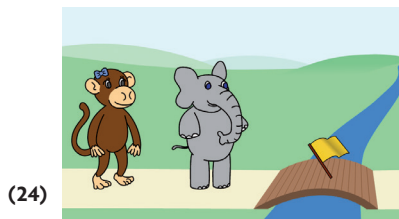
(17)



Toen begon het te regenen. Gelukkig was er een grot dichtbij. Anita Aap en Oscar Olifant liepen snel de donkere grot binnen en schuilden voor de regen. (12) Toen hun ogen gewend waren aan het donker, zagen ze iets in de grot liggen. [TS+Q] (13) Oscar en Anita liepen verder de grot in en ontdekten nog iets op de grond. [TS+Q] (14) Achter een grote steen lag ook nog iets verstopt. [TS+Q] Toen stopte het met regenen. Anita Aap en Oscar Olifant liepen de grot uit en gingen weer verder. Ze kwamen aan bij een weiland. Tussen het gras groeiden bloemetjes in alle kleuren van de regenboog. (15) Opeens zagen ze iets tussen de bloemetjes liggen. [TS+Q] Oscar en Anita liepen verder door het weiland. Ze stopten zo nu en dan om aan de bloemetjes te ruiken. (16) Toen zagen ze nog iets tussen de bloemetjes liggen. [TS+Q] (17) Verderop achter een grote zonebloem lag ook nog iets. [TS+Q]



Anita Aap en Oscar Olifant liepen weer verder. Na een tijdje kwamen ze bij een vijver. (18) Ze zagen iets in de vijver liggen. [TS+Q] (19) Ze vonden ook nog iets tussen het riet. [TS+Q] Het water van de vijver was lekker warm. Oscar en Anita sprongen in het water en spetterden elkaar nat. (20) Ineens zagen ze iets in het water drijven. [TS+Q] Anita Aap en Oscar Olifant gingen weer uit het water en lieten zich opdrogen in de zon. Daarna liepen ze weer verder. (21) Opeens zagen ze iets voor een rots liggen. [TS+Q] (22) Ze klommen op de rots en zagen op de top nog iets liggen. [TS+Q] Bovenop de rots waaide een frisse wind. Ze hadden een prachtig uitzicht. Oscar en Anita gingen aan de andere kant van de rots naar beneden. (23) Daar lag ook iets. [TS+Q]



Anita en Oscar liepen weer verder. Ze kwamen bij een beekje en liepen naar de brug. (24) Er lag iets op de brug. [TS+Q] (25) Aan de andere kant van de brug lag ook iets. [TS+Q] Ineens hoorden Oscar en Anita een vreemd geluid. Wat was dat? Ze liepen verder in de richting waar het geluid vandaan kwam. (26) Daar zat een hond tegen een boom te slapen. De hond snurkte heel hard. Vlak naast de hond lag weer iets in het gras. [TS+Q] Dat was wel heel toevallig. Overal kwamen ze spulletjes tegen die er precies zo uitzagen als hun eigen spulletjes. Opeens hadden Anita en Oscar het door. Het waren hun eigen spulletjes! Maar hoe kon dat? Had de hond hun spulletjes gestolen? (27) Ze maakten de hond wakker en vroegen hem: “Heb jij onze spulletjes gestolen?”. Hond werd vuurrood en zei: “Ja, ik heb jullie spulletjes meegenomen. Het spijt me heel erg. Ik wilde alleen maar spelen.” Anita en Oscar dachten even na. “Wij zijn niet boos, hoor”, zei Anita. “Maar dan moet je wel al onze spulletjes weer terug brengen”, zei Oscar. Hond zei: “Ik zal alles meteen weer terugbrengen.” Hond ging gelijk op pad. Hij zocht alle spulletjes bij elkaar en bracht ze naar de tuin van Anita en Oscar. (28) Maar nu heeft Hond een probleem! Welke dingen zijn van Oscar Olifant en welke van Anita Aap? Alles ligt door elkaar. Kun jij Hond helpen? Het geeft niet als je niet alles meer weet hoor. (29) Van wie is de auto?

Translation of the narrative

Explanatory note: The appearance of a certain picture is indicated in round brackets (e.g., (1)). The pictures can be found above, next to the Dutch narrative. '[TS+Q]' indicates that a test sentence is presented, followed by the respective question (see Tables C1 and C2).

(1) These are Anita Aap and Oscar Olifant. Anita Aap is a girl and Oscar Olifant is a boy. Who is Anita Aap? Who is Oscar Olifant? Who is a girl? Who is a boy? (2) Oscar Olifant and Anita Aap are best friends. They live next to each other in two houses. Oscar Olifant lives in a stone house and Anita Aap lives in a wooden house. They see each other every day and experience nice adventures together. (1) One day, Anita Aap and Oscar Olifant woke up early. As always, they started the day with morning gymnastics. Oscar and Anita were both doing something: She/He did a handstand. Who did a handstand? She/He ran around the houses. Who ran around the houses? She/He skipped rope. Who skipped rope? She/He did a somersault. Who did a somersault? After exercising, Oscar Olifant and Anita Aap decided to go for a walk together. It was a beautiful day and already nice and warm in the morning. They walked bare-foot through the grass. (3) Then they saw something lying in the grass. [TS+Q] (4) There was something else lying in the grass as well. [TS+Q] (5) Anita and Oscar walked on and discovered yet another thing. [TS+Q] Oscar and Anita continued their walk. They reached a path and followed it for a while. (6) Suddenly they saw something lying on the path. [TS+Q] (7) And there was also something lying *next* to the path. [TS+Q] Anita and Oscar walked a bit further. (8) At the crossroads, another thing was lying on the ground. [TS+Q] They continued walking. Anita Aap whistled a song and Oscar Olifant trumpeted with his trunk. They had a great time together. (9) Suddenly, they saw something lying under a tree. [TS+Q] (10) Oscar and Anita looked up and discovered that there was also something hanging in the tree. [TS+Q] (11) They walked around the tree and saw on the other side also something hanging in the tree. [TS+Q] Then it started raining. Luckily, a cave was close by. Anita Aap and Oscar Olifant quickly entered the dark cave to shelter from the rain. (12) When their eyes were used to the darkness, they saw something lying in the cave. [TS+Q] (13) Oscar and Anita walked deeper into the cave and discovered yet another thing on the floor. [TS+Q] (14) Behind a big rock was another object hidden. [TS+Q] Then it stopped raining. Anita Aap and Oscar Olifant exited the cave and continued their walk. They came to a meadow. Between the grass, flowers in all colors of the rainbow were growing. (15) Suddenly they discovered something between the flowers. [TS+Q] Oscar and Anita continued walking through the meadow. They stopped now and then to smell the flowers. (16) Then they found another thing between the flowers. [TS+Q] (17) A bit further behind a big sunflower, there was yet another object. [TS+Q] Anita Aap and Oscar Olifant continued their walk. After a while, they arrived at a pond. (18) They saw something in the pond. [TS+Q] (19) They found also something among the reeds. [TS+Q] The water of the pond was nice and warm. Oscar and Anita jumped into the water and splashed each other. (20) Suddenly they saw

something floating in the water. [TS+Q] Anita Aap and Oscar Olifant went out of the water and dried themselves in the sun. Afterwards they walked on. (21) Suddenly they saw something lying in front of a rock. [TS+Q] (22) They climbed the rock and found another thing on top. [TS+Q] On top of the rock was a fresh breeze. They had a wonderful view. Oscar and Anita went down the rock on the other side. (23) There was also something. [TS+Q] Anita and Oscar walked on. They reached a little stream and walked to the bridge. (24) There was something lying on the bridge. [TS+Q] (25) On the other side of the bridge was yet another thing. [TS+Q] Suddenly, Oscar and Anita heard a strange sound. What was that? They walked into the direction of the sound. (26) There was a dog leaning against a tree, sleeping. The dog was snoring very loudly. Close to the dog was yet again something lying in the grass [TS+Q]. That was a strange coincidence! Everywhere they encountered things that looked exactly like their own things. Suddenly Anita and Oscar realized: These WERE their things! But how was this possible? Did the dog steal their things? They woke the dog and asked him: "Did you steal our things?" Dog turned as red as fire and said: "Yes, I took your things. I'm very sorry. I just wanted to play." Anita and Oscar thought about it for a while. "We are not angry", said Anita. "But then you have to bring all our things back", said Oscar. Dog said: "I promise to bring everything back immediately." Dog went on his way directly. He gathered up all things and brought them to the garden of Anita and Oscar. (28) But now Dog has a problem! Which things belong to Oscar Olifant and which ones to Anita Aap? Everything is mixed up. Can you help Dog? Don't worry if you don't remember everything. (29) To whom does the car belong?

REFERENCES

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Anand, P. (2006). *De de se*. Ph.D. thesis, MIT.
- Anand, P., & Nevins, A. (2004). Shifty operators in changing contexts. *Proceedings of SALT XIV*, 20–37. <http://doi.org/10.3765/salt.v14i0.2913>
- Andeweg, S., Hendrix, R., Hoff, V. van't, & Hoop, H. de. (2013). En dan vermoord je hem ... Over de invloed van grammaticale persoon op identificatie. *Neerlandistiek.nl*, 13(1), 1–19.
- Arnold, J. E., Brown-Schmidt, S., & Trueswell, J. (2007). Children's use of gender and order-of-mention during pronoun comprehension. *Language and Cognitive Processes*, 22(4), 527–565. <http://doi.org/10.1080/01690960600845950>
- Arslan, B., & Verbrugge, R. (2015). Teaching children to attribute second-order false beliefs: A training study with feedback. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*. Austin, TX, 1, 108–113.
- Audring, J. (2009). *Reinventing Pronoun Gender*. Doctoral dissertation, University of Utrecht.
- Auer, P. (1998). Zwischen Parataxe und Hypotaxe: "abhängige Hauptsätze" im gesprochenen und geschriebenen Deutsch. *Zeitschrift Für Germanistische Linguistik*, 26(3), 284–307. <http://doi.org/10.1515/zfgl.1998.26.3.284>
- Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics*. Cambridge: Cambridge University Press. <http://doi.org/10.1017/CBO9780511801686>
- Baker, C. D., & Freebody, P. (1989). *Children's First School Books*. Oxford: Basil Blackwell.
- Banfield, A. (1973). Narrative style and the grammar of direct and indirect speech. *Foundations of Language*, 10(1), 1–39.
- Banfield, A. (1982). *Unspeakable Sentences: Narration and Representation in the Language of Fiction*. Boston, London, Melbourne, Henley: Routledge & Kegan Paul.
- Barbieri, F., & Eckhardt, S. E. B. (2007). Applying corpus-based findings to form-focused instruction: The case of reported speech. *Language Teaching Research*, 11(3), 319–346. <http://doi.org/10.1177/1362168807077563>
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, 21(1), 37–46. [http://doi.org/10.1016/0010-0277\(85\)90022-8](http://doi.org/10.1016/0010-0277(85)90022-8)
- Barsalou, L. W. (1999). Language comprehension: Archival memory or preparation for situated action? *Discourse Processes*, 28(1), 61–80. <http://doi.org/10.1080/01638539909545069>
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645. <http://doi.org/10.1146/annurev.psych.59.103006.093639>
- Bates, E. (1990). Language about me and you: Pronominal reference and the emerging concept of self. In D. Cicchetti & M. Beeghly (Eds.), *The Self in Transition. Infancy to Childhood* (pp. 165–182). Chicago, London: University of Chicago Press.
- Bauman, R. (1986). *Story, Performance, and Event: Contextual Studies of Oral Narrative*. Cambridge: Cambridge University Press.

- Baynham, M. (1996). Direct speech. What's it doing in non-narrative discourse? *Journal of Pragmatics*, 25, 61–81. [http://doi.org/10.1016/0378-2166\(94\)00074-3](http://doi.org/10.1016/0378-2166(94)00074-3)
- Berman, R. A., & Slobin, D. I. (Eds.). (1994). *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Bloom, L., Lightbown, P., Hood, L., Bowerman, M., Maratsos, M., & Maratsos, M. P. (1975). Structure and variation in child language. *Monographs of the Society for Research in Child Development*, 40(2), 1–97.
- Boeg Thomsen, D. (2014). Anchoring perspectives: Danish children's linguistic strategies for managing representations of reality. *Paper Presented at the Conference "(R)epresenting the Speech of Others", Groningen, 2014*.
- Bohan, J., Sanford, A. J., Cochrane, S., & Sanford, A. J. S. (2008). Direct and indirect speech modulates depth of processing. *Paper presented at the 14th Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP), Cambridge, UK*.
- Bol, G. W., & Kasparian, K. (2009). The production of pronouns in Dutch children with developmental language disorders: A comparison between children with SLI, hearing impairment, and Down's syndrome. *Clinical Linguistics & Phonetics*, 23(9), 631–646. <http://doi.org/10.1080/02699200902995677>
- Bol, G. W., & Kuiken, F. (1986). Het gebruik van pronomina bij kinderen van een tot vier. *Toegepaste Taalwetenschap in Artikelen*, 24(1), 47–68. <http://doi.org/10.1075/ttwia.24.06bol>
- Bol, G. W., & Kuiken, F. (1988). *Grammaticale Analyse van Taalontwikkelingsstoornissen*. Doctoral dissertation, University of Amsterdam.
- Boogart, R. (1996). Tense and temporal ordering in English and Dutch indirect speech. In T. A. J. M. Janssen & W. van der Wurff (Eds.), *Reported Speech: Forms and Functions of the Verb* (pp. 213–235). Amsterdam, Philadelphia: John Benjamins.
- Booij, G. (2002). *The Morphology of Dutch*. Oxford: Oxford University Press.
- Borghì, A. M., Glenberg, A. M., & Kaschak, M. P. (2004). Putting words in perspective. *Memory & Cognition*, 32(6), 863–873. <http://doi.org/10.3758/BF03196865>
- Bortolussi, M., & Dixon, P. (2003). *Psychonarratology. Foundations for the Empirical Study of Literary Response*. Cambridge: Cambridge University Press.
- Bosch, P., Katz, G., & Umbach, C. (2007). The non-subject bias of German demonstrative pronouns. In M. Schwarz-Friesel, M. Consten, & M. Knees (Eds.), *Anaphors in Text: Cognitive, Formal and Applied Approaches to Anaphoric Reference* (pp. 145–164). Amsterdam, Philadelphia: John Benjamins.
- Bosch, P., Rozario, T., & Zhao, Y. (2003). Demonstrative pronouns and personal pronouns: German “der” vs. “er.” *Proceedings of the EACL 2003, Workshop on the Computational Treatment of Anaphora*.
- Bosch, P., & Umbach, C. (2007). Reference determination for demonstrative pronouns. *ZAS Papers in Linguistics*, (48), 39–51.
- 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. <http://doi.org/10.1353/lan.2010.0010>
- Bredel, U. (2004). Die Didaktik der Gänsefüßchen. In U. Bredel, G. Siebert-Ott, & T. Thelen (Eds.), *Schriftspracherwerb und Orthographie* (pp. 207–240). Schneider Hohengehren: Baltmannsweiler.

- Brener, R. (1983). Learning the deictic meaning of third person pronouns. *Journal of Psycholinguistic Research*, 12(3), 235–262. <http://doi.org/10.1007/BF01067669>
- Brocki, K. C., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Developmental Neuropsychology*, 26(2), 571–593. http://doi.org/10.1207/s15326942dn2602_3
- Brunyé, T. T., Ditman, T., Mahoney, C. R., Augustyn, J. S., & Taylor, H. A. (2009). When you and I share perspectives: Pronouns modulate perspective taking during narrative comprehension. *Psychological Science*, 20(1), 27–32. <http://doi.org/10.1111/j.1467-9280.2008.02249.x>
- Brunyé, T. T., Ditman, T., Mahoney, C. R., & Taylor, H. a. (2011). Better you than I: Perspectives and emotion simulation during narrative comprehension. *Journal of Cognitive Psychology*, 23(5), 659–666. <http://doi.org/10.1080/20445911.2011.559160>
- Bühler, K. (1934). *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Jena: Gustav Fischer.
- Burke, D. M., & Shafto, M. A. (2008). Language and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The Handbook of Aging and Cognition. 3rd edition* (Vol. 11, pp. 3–60). New York: Psychology Press. <http://doi.org/10.1097/00011363-198109000-00010>
- Cappelen, H., & Lepore, E. (1997). Varieties of quotation. *Mind*, 106, 429–450. <http://doi.org/10.1075/bjl.17.04cap>
- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, 72(4), 1032–1053. <http://doi.org/10.1111/1467-8624.00333>
- Carter, C. A., Bottoms, B. L., & Levine, M. (1996). Linguistic and socioemotional influences on the accuracy of children's reports. *Law and Human Behavior*, 20(3), 335–358. <http://doi.org/10.1007/BF01499044>
- Chafe, W. L. (1982). Integration and involvement in speaking, writing, and oral literature. In D. Tannen (Ed.), *Spoken and Written Language: Exploring Orality and Literacy* (pp. 35–53). Norwood, New Jersey: Ablex Publishing Corporation.
- Charney, R. (1980). Speech roles and the development of personal pronouns. *Journal of Child Language*, 7(3), 509–528. <http://doi.org/10.1017/S0305000900002816>
- Chiat, S. (1981). Context-specificity and generalization in the acquisition of pronominal distinctions. *Journal of Child Language*, 8(1), 75–91. <http://doi.org/10.1017/S0305000900003020>
- Chiat, S. (1982). If I were you and you were me: The analysis of pronouns in a pronoun-reversing child. *Journal of Child Language*, 9(2), 359–379. <http://doi.org/10.1017/S0305000900004761>
- Chiat, S. (1986). Personal pronouns. In P. Fletcher & M. Garman (Eds.), *Language Acquisition. Studies in First Language Development* (pp. 339–355). Cambridge: Cambridge University Press.
- Chien, Y.-C., & Wexler, K. (1990). Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition*, 1(3), 225–295. http://doi.org/10.1207/s15327817la0103_2
- Clark, H. H., & Gerrig, R. J. (1990). Quotations as demonstrations. *Language*, 66(4), 764–805. <http://doi.org/10.2307/414729>
- Clements, W. A., Rustin, C. L., & McCallum, S. (2000). Promoting the transition from implicit to explicit understanding: A training study of false belief. *Developmental Science*, 3(1), 81–92.

- Comrie, B. (1997). Pragmatic binding: Demonstratives as anaphors in Dutch. *Proceedings of the Twenty-Third Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Pragmatics and Grammatical Structure*, 50–61.
- Cordeiro, P., Giacobbe, M. E., & Cazden, C. (1983). Apostrophes, quotation marks, and periods: Learning punctuation in the first grade. *Language Arts*, 60(3), 323–332.
- Coulmas, F. (1985). Direct and indirect speech: General problems and problems of Japanese. *Journal of Pragmatics*, 9, 41–63. [http://doi.org/10.1016/0378-2166\(85\)90047-5](http://doi.org/10.1016/0378-2166(85)90047-5)
- Coulmas, F. (1986). Reported speech: Some general issues. In F. Coulmas (Ed.), *Direct and Indirect Speech* (pp. 1–28). Berlin, New York, Amsterdam: Mouton de Gruyter.
- Dale, P. S., & Crain-Thoreson, C. (1993). Pronoun reversals: Who, when, and why? *Journal of Child Language*, 20(3), 573–589. <http://doi.org/10.1017/S0305000900008485>
- Davidson, D. (1979). Quotation. *Theory and Decision*, 11, 27–40.
- Davidson, K. (2015). Quotation, demonstration, and iconicity. *Linguistics & Philosophy*, 38(6), 477–520.
- De Villiers, J. G. (2005). 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: Oxford University Press.
- De Villiers, J., Hobbs, K., & Hollebrandse, B. (2014). Recursive complements and propositional attitudes. In T. Roeper & M. Speas (Eds.), *Recursion: Complexity in Cognition* (pp. 221–242). Cham: Springer. <http://doi.org/10.1007/978-3-319-05086-7>
- De Villiers, J. G., & De Villiers, P. A. (2000). Linguistic determinism and the understanding of false beliefs. In P. Mitchell & K. J. Riggs (Eds.), *Children's Reasoning and the Mind* (pp. 191–228). Hove: Psychology Press.
- 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, 1037–1060. [http://doi.org/10.1016/S0885-2014\(02\)00073-4](http://doi.org/10.1016/S0885-2014(02)00073-4)
- De Villiers, P. A., & De Villiers, J. G. (2012). Deception dissociates from false belief reasoning in deaf children: Implications for the implicit versus explicit theory of mind distinction. *The British Journal of Developmental Psychology*, 30, 188–209. <http://doi.org/10.1111/j.2044-835X.2011.02072.x>
- Deutsch, W., & Pechmann, T. (1978). Ihr, dir, or mir? On the acquisition of pronouns in German children. *Cognition*, 6, 155–168. [http://doi.org/10.1016/0010-0277\(78\)90020-3](http://doi.org/10.1016/0010-0277(78)90020-3)
- Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition*, 11(2), 159–184. [http://doi.org/10.1016/0010-0277\(82\)90024-5](http://doi.org/10.1016/0010-0277(82)90024-5)
- Diessel, H. (2004). *The Acquisition of Complex Sentences*. Cambridge: Cambridge University Press.
- Diessel, H. (2012). Deixis and demonstratives. In C. Maienborn, K. von Stechow, & P. Portner (Eds.), *Semantics (HSK 33.3)* (Vol. III, pp. 1–25). Berlin, Boston: De Gruyter.
- Diessel, H., & Tomasello, M. (2001). The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics*, 12(2), 97–141. <http://doi.org/10.1515/cogl.12.2.97>
- Ditman, T., Bruny , T. T., Mahoney, C. R., & Taylor, H. A. (2010). Simulating an enactment effect: Pronouns guide action simulation during narrative comprehension. *Cognition*, 115(1), 172–178. <http://doi.org/10.1016/j.cognition.2009.10.014>
- Duursma, E. (2014). Parental bookreading practices among families in the Netherlands. *Journal of Early Childhood Literacy*, 14(4), 435–458. <http://doi.org/10.1177/1468798414523025>

- Ebert, K. (1986). Reported speech in some languages of Nepal. In F. Coulmas (Ed.), *Direct and Indirect Speech* (pp. 145–159). Berlin, New York, Amsterdam: Mouton de Gruyter.
- Eckardt, R. (2015). *The Semantics of Free Indirect Discourse: How Texts allow us to Mind-Read and Eavesdrop*. Leiden: Koninklijke Brill NV.
- Eerland, A., Engelen, J. A. A., & Zwaan, R. A. (2013). The influence of direct and indirect speech on mental representations. *PLOS ONE*, 8(6), 1–9. <http://doi.org/10.1371/journal.pone.0065480>
- Elbourne, P. (2005). On the acquisition of principle B. *Linguistic Inquiry*, 36(3), 333–365. <http://doi.org/10.1162/0024389054396908>
- Ely, R., & McCabe, A. (1993). Remembered voices. *Journal of Child Language*, 20(3), 671–696. <http://doi.org/10.1017/S0305000900008539>
- Emmorey, K., & Reilly, J. (1998). The development of quotation and reported action: Conveying perspective in ASL. In E. V. Clark (Ed.), *The Proceedings of the Twenty-ninth Annual Child Language Research Forum* (pp. 81–90).
- Engberg-Pedersen, E. (1993). *Space in Danish Sign Language: The Semantics and Morphosyntax of the Use of Space in Visual Language*. Hamburg: Signum.
- Engberg-Pedersen, E. (1995). Point of view expressed through shifters. In K. Emmorey & J. S. Reilly (Eds.), *Language, Gesture, and Space* (pp. 133–154). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Engberg-Pedersen, E., & Christensen, R. V. (submitted). Mental states and activities in Danish narratives: Children with autism and children with language impairment.
- Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, 40, 760–768. <http://doi.org/10.1016/j.jesp.2004.02.002>
- Evans, K. E., & Demuth, K. (2012). Individual differences in pronoun reversal: Evidence from two longitudinal case studies. *Journal of Child Language*, 39(1), 162–91. <http://doi.org/10.1017/S0305000911000043>
- Evans, N. (2012). Some problems in the typology of quotation: A canonical approach. In D. Brown, M. Chumakina, & G. G. Corbett (Eds.), *Canonical Morphology and Syntax* (pp. 66–98). Oxford: Oxford University Press. <http://doi.org/10.1093/acprof:oso/9780199604326.003.0004>
- Fabricius-Hansen, C. (2002). Nicht-direktes Referat im Deutschen. Typologie und Abgrenzungsprobleme. In C. Fabricius-Hansen, O. Leirbukt, & O. Letnes (Eds.), *Modus, Modalverben, Modalpartikeln* (pp. 7–29). Trier: Wissenschaftlicher Verlag.
- Ferreiro, E., & Pontecorvo, C. (1999). Managing the written text: The beginning of punctuation in children's writing. *Learning and Instruction*, 9(6), 543–564. [http://doi.org/10.1016/S0959-4752\(99\)00006-7](http://doi.org/10.1016/S0959-4752(99)00006-7)
- Ferreiro, E., & Zuccheromaglio, C. (1996). Children's use of punctuation marks: The case of quoted speech. In C. Pontecorvo, M. Orsolini, B. Burge, & L. B. Resnick (Eds.), *Children's Early Text Construction* (pp. 177–205). Mahwah, NJ: Lawrence Erlbaum Associates.
- Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the appearance-reality distinction. *Cognitive Psychology*, 15(1), 95–120. [http://doi.org/10.1016/0010-0285\(83\)90005-1](http://doi.org/10.1016/0010-0285(83)90005-1)
- Flobbe, L., Verbrugge, R., Hendriks, P., & Krämer, I. (2008). Children's application of theory of mind in reasoning and language. *Journal of Logic, Language and Information*, 17(4), 417–442. <http://doi.org/10.1007/s10849-008-9064-7>

- Fludernik, M. (1993). *The Fictions of Language and the Languages of Fiction: The Linguistic Representation of Speech and Consciousness*. London, New York: Routledge.
- Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift Für Philosophie Und Philosophische Kritik*, 100(1), 25–50.
- Frith, U., & De Vignemont, F. (2005). Egocentrism, allocentrism, and Asperger syndrome. *Consciousness and Cognition*, 14(4), 719–738. <http://doi.org/10.1016/j.concog.2005.04.006>
- Galatolo, R. (2007). Active voicing in court. In E. Holt & R. Clift (Eds.), *Reporting Talk: Reported Speech in Interaction* (pp. 195–220). Cambridge: Cambridge University Press.
- Gernsbacher, M. a., & Faust, M. E. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(2), 245–262. <http://doi.org/10.1037/0278-7393.17.2.245>
- Gerrig, R. J. (1993). *Experiencing Narrative Worlds: On the Psychological Activities of Reading*. New Haven, London: Yale University Press.
- Geurts, B., & Maier, E. (2005). Quotation in context. *Belgian Journal of Linguistics*, 17(1), 109–128. <http://doi.org/10.1075/bjl.17.07geu>
- Goodell, E. W., & Sachs, J. (1992). Direct and indirect speech in English-speaking children's retold narratives. *Discourse Processes*, 15(4), 395–422. <http://doi.org/10.1080/01638539209544820>
- 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(1), 26–37. <http://doi.org/10.1111/j.1467-8624.1988.tb03192.x>
- Groenewold, R. (2015). *Direct and Indirect Speech in Aphasia: Studies of Spoken Discourse Production and Comprehension*. Doctoral dissertation, University of Groningen.
- Groenewold, R., Bastiaanse, R., & Huiskes, M. (2013). Direct speech constructions in aphasic Dutch narratives. *Aphasiology*, 27(5), 546–567. <http://doi.org/10.1080/02687038.2012.742484>
- Groenewold, R., Bastiaanse, R., Nickels, L., & Huiskes, M. (2014). Perceived liveliness and speech comprehensibility in aphasia: The effects of direct speech in auditory narratives. *International Journal of Language & Communication Disorders*, 49(4), 486–497. <http://doi.org/10.1111/1460-6984.12080>
- Groenewold, R., Bastiaanse, R., Nickels, L., Wieling, M., & Huiskes, M. (2014). The effects of direct and indirect speech on discourse comprehension in Dutch listeners with and without aphasia. *Aphasiology*, 28(7), 862–884. <http://doi.org/10.1080/02687038.2014.902916>
- Günthner, S. (1997). The contextualization of affect in reported dialogues. In S. Niemeier & R. Dirven (Eds.), *The Language of Emotions. Conceptualization, Expression, and Theoretical Foundation* (pp. 247–276). Amsterdam: John Benjamins.
- Haan, G. J. de (2010). *Studies in West Frisian Grammar*. (J. Hoekstra, W. Visser, & G. Jensma, Eds.). Amsterdam, Philadelphia: John Benjamins.
- Haberland, H. (1986). Reported speech in Danish. In F. Coulmas (Ed.), *Direct and Indirect Speech* (pp. 219–248). Berlin: De Gruyter.
- Hale, C. M., & Tager-Flusberg, H. (2003). The influence of language on theory of mind: A training study. *Developmental Science*, 6(3), 346–59. <http://doi.org/10.1111/1467-7687.00289>
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

- Harris, J. A., & Potts, C. (2009). Perspective-shifting with appositives and expressives. *Linguistics and Philosophy*, 32(6), 523–552. <http://doi.org/10.1007/s10988-010-9070-5>
- Harris, P. L. (2000). *The Work of the Imagination*. Oxford: Blackwell.
- Heim, I. (1991). Artikel und Definitheit. In H. Steger & H. E. Wiegand (Eds.), *Handbücher zur Sprach- und Kommunikationswissenschaft, Bd. 6 Semantik* (pp. 487–535). Berlin, New York: Walter de Gruyter.
- Hendriks, P. (2014). *Asymmetries between Language Production and Comprehension*. Dordrecht: Springer.
- Hendriks, P., Koster, C., & Hoeks, J. C. J. (2014). Referential choice across the lifespan: Why children and elderly adults produce ambiguous pronouns. *Language, Cognition and Neuroscience*, 29(4), 391–407. <http://doi.org/10.1080/01690965.2013.766356>
- Hendriks, P., & Spenader, J. (2006). When production precedes comprehension: An optimization approach to the acquisition of pronouns. *Language Acquisition*, 13(4), 319–348. <http://doi.org/10.1207/s15327817la1304>
- Hendriks, P., van Rijn, H., & Valkenier, B. (2007). Learning to reason about speakers' alternatives in sentence comprehension: A computational account. *Lingua*, 117, 1879–1896. <http://doi.org/10.1016/j.lingua.2006.11.008>
- Herrmann, A., & Steinbach, M. (2007). Wenn "ich" nicht ich ist: Zitieren in Gebärdensprachen. In E. Brendel, J. Meibauer, & M. Steinbach (Eds.), *Zitat und Bedeutung. Linguistische Berichte, Sonderheft 15* (pp. 153–179). Hamburg: Helmut Buske Verlag.
- Herrmann, A., & Steinbach, M. (2012). Quotation in sign languages: A visible context shift. In I. Buchstaller & I. Van Alphen (Eds.), *Quotatives. Cross-linguistic and Cross-disciplinary Perspectives* (pp. 203–228). Amsterdam: John Benjamins.
- Heusinger, K. von (2002). Reference and representation of pronouns. In H. J. Simon & H. Wiese (Eds.), *Pronouns—Grammar and Representation* (pp. 109–135). Amsterdam, Philadelphia: John Benjamins.
- Hewlett, N., Kelsey, C., & Lickley, R. (2003). Children's perception of direct and indirect reported speech. *Proceedings of 15th International Congress of Phonetic Sciences (ICPhS)*, 1313–1316.
- Hickmann, M. (1993). The boundaries of reported speech in narrative discourse: Some developmental aspects. In J. A. Lucy (Ed.), *Reflexive Language: Reported Speech and Metapragmatics* (pp. 63–90). Cambridge: Cambridge University Press.
- Hickmann, M. (1995). Discourse organization and the development of reference to person, space, and time. In P. Fletcher & B. MacWhinney (Eds.), *The Handbook of Child Language* (pp. 194–218). Oxford: Blackwell.
- Hobson, R. P., Lee, A., & Hobson, J. A. (2010). Personal pronouns and communicative engagement in autism. *Journal of Autism and Developmental Disorders*, 40(6), 653–664. <http://doi.org/10.1007/s10803-009-0910-5>
- Hollebrandse, B. (2007). A special case of wh-extraction in child language. *Lingua*, 117(11), 1897–1906. <http://doi.org/10.1016/j.lingua.2006.09.006>
- Hoop, H. de, & Tarenskeen, S. (2014). It's all about you in Dutch. *Journal of Pragmatics*, 1–36. <http://doi.org/10.1016/j.pragma.2014.07.001>
- Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, 50(3), 346–363. <http://doi.org/10.1002/bimj.200810425>
- Houwen, F. van der (2012). The effect of genre on reporting speech: Conversations and newspaper articles. *Borealis: An International Journal of Hispanic Linguistics*, 1(1), 101–130. <http://doi.org/10.7557/1.1.1.2294>

- Hübl, A. (2013). Context shift (im)possible: Indexicals in German Sign Language. *Proceedings of ConSOLE XXI*, 171–183.
- Huizinga, M., Dolan, C. V., & Van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, 44(11), 2017–2036. <http://doi.org/10.1016/j.neuropsychologia.2006.01.010>
- Hunter, J. (2014). Structured contexts and anaphoric dependencies. *Philosophical Studies*, 168(1), 35–58. <http://doi.org/10.1007/s11098-013-0209-4>
- Imbens-Bailey, A., & Pan, A. (1998). The pragmatics of self- and other-reference in young children. *Social Development*, 7(2), 219–233. <http://doi.org/10.1111/1467-9507.00063>
- Jansen, W., Gregory, M. L., & Brenier, J. M. (2001). Prosodic correlates of directly reported speech: Evidence from conversational speech. *Proceedings of the ISCA Workshop on Prosody in Speech Recognition and Understanding* (pp. 77–80). Red Bank, NJ.
- Jasnow, M., & Feldstein, S. (1986). Adult-like temporal characteristics of mother-infant vocal interactions. *Child Development*, 57(3), 754–761. <http://doi.org/10.2307/1130352>
- Johnson, M., & Lepore, E. (2011). Misrepresenting misrepresentation. In E. Brendel, J. Meibauer, & M. Steinbach (Eds.), *Understanding Quotation* (pp. 231–248). Berlin, New York: De Gruyter Mouton.
- Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge: Cambridge University Press.
- Kalmanovitch, Y. (2015). Jumping out of context—jumping out of tone. *Paper Presented at the International Conference of Phonetic Sciences, Glasgow*.
- Kaplan, D. (1989). Demonstratives. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 481–563). Oxford: Oxford University Press.
- Karmiloff-Smith, A. (1985). Language and cognitive processes from a developmental perspective. *Language and Cognitive Processes*, 1(1), 61–85. <http://doi.org/10.1080/01690968508402071>
- Kasimir, E. (2008). Prosodic correlates of subclausal quotation marks. *ZAS Papers in Linguistics*, 49, 67–77.
- Kim, S. H., Paul, R., Tager-Flusberg, H., & Lord, C. (2014). Language and communication in autism. In F. R. Volkmar, S. J. Rogers, R. Paul, & K. A. Pelphrey (Eds.), *Handbook of Autism and Pervasive Developmental Disorders* (4th ed., pp. 230–262). New York: John Wiley & Sons.
- Klewitz, G., & Couper-Kuhlen, E. (1999). Quote-unquote? The role of prosody in the contextualization of reported speech sequences. *Interaction and Linguistic Structures*, 12, 1–34.
- Köder, F. (to appear). Reporting vs. pretending. Degrees of identification in role play and reported speech. In A. Hübl & M. Steinbach (Eds.), *Linguistic Foundations of Narration in Spoken and Sign Languages*. John Benjamins.
- Köder, F. (2013). How children acquire reported speech in German and Dutch: A corpus study. In B. Sonnenhauser, C. Trautmann, & P. Noel (Eds.), „*Perspektiven*“. *Diskussionsforum Linguistik in Bayern / Bavarian Working Papers in Linguistics 2* (pp. 15–28). München, Bamberg.
- Köder, F., & Maier, E. (2015). Children mix direct and indirect speech: Evidence from pronoun comprehension. *Journal of Child Language*. <http://doi.org/10.1017/S0305000915000318>

- Köder, F., Maier, E., & Hendriks, P. (2015). Perspective shift increases processing effort of pronouns: A comparison between direct and indirect speech. *Language, Cognition and Neuroscience*, 30(8), 940–946. <http://doi.org/10.1080/23273798.2015.1047460>
- Koev, T. (2015). Quotational Indefinites. MS. Retrieved from <http://todorkoev.weebly.com/papers.html>
- Koev, T. (2013). *Apposition and the Structure of Discourse*. Ph.D. thesis, Rutgers University.
- Koster, C. (1993). *Errors in Anaphora Acquisition*. Doctoral dissertation, Utrecht University.
- Koster, C., Hoeks, J., & Hendriks, P. (2011). Comprehension and production of subject pronouns: Evidence for the asymmetry of grammar. In A. Grimm, A. Müller, C. Hamann, & E. Ruigendijk (Eds.), *Production-Comprehension Asymmetries in Child Language* (pp. 99–122). Berlin: De Gruyter Mouton.
- Kratzer, A. (1999). Beyond “Ouch” and “Oops.” How descriptive and expressive meaning interact. *Cornell Conference on Theories of Context Dependence*.
- Lee, A., Hobson, R. P., & Chiat, S. (1994). I, you, me, and autism: An experimental study. *Journal of Autism and Developmental Disorders*, 24(2), 155–176. <http://doi.org/10.1007/BF02172094>
- Leech, G. N., & Short, M. H. (1981). *Style in Fiction: A Linguistic Introduction to English Fictional Prose*. London, New York: Longman.
- Legendre, G., & Smolensky, P. (2012). On the asymmetrical difficulty of acquiring person reference in French: Production versus comprehension. *Journal of Logic, Language and Information*, 21(1), 7–30. <http://doi.org/10.1007/s10849-011-9150-0>
- Lehrer, A. (1989). Remembering and representing prose: Quoted speech as a data source. *Discourse Processes*, 12(1), 105–125. <http://doi.org/10.1080/01638538909544721>
- Levinson, S. C. (2004). Deixis. In L. R. Horn & G. Ward (Eds.), *The Handbook of Pragmatics* (pp. 97–121). Oxford: Blackwell.
- Lewis, D. (1979). Attitudes de dicto and de se. *The Philosophical Review*, 88(4), 513–543.
- Li, C. (1986). Direct and indirect speech: A functional study. In F. Coulmas (Ed.), *Direct and Indirect Speech* (pp. 29–45). Berlin, New York, Amsterdam: Mouton de Gruyter.
- Lillo-Martin, D., & de Quadros, R. M. (2011). Acquisition of the syntax-discourse interface: The expression of point of view. *Lingua*, 121(4), 623–636. <http://doi.org/10.1016/j.lingua.2010.07.001>
- 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. <http://doi.org/10.1111/1467-8624.00597>
- Loveland, K. A. (1984). Learning about points of view: Spatial perspective and the acquisition of “I/you”. *Journal of Child Language*, 11(3), 535–56. <http://doi.org/10.1017/S0305000900005948>
- Lyons, J. (1977). *Semantics* (Vol. 2). Cambridge: Cambridge University Press. <http://doi.org/10.1017/CBO9780511620614>
- MacWhinney, B. (2000). *The CHILDES Project. Tools for Analyzing Talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Maier, E. (2009). Japanese reported speech: Against a direct-indirect distinction. In H. Hattori, T. Kawamura, T. Idé, M. Yokoo, & Y. Murakami (Eds.), *New Frontiers in Artificial Intelligence* (Vol. 5447, pp. 133–145). Berlin, Heidelberg: Springer. http://doi.org/10.1007/978-3-642-00609-8_13
- Maier, E. (2014). Mixed quotation: The grammar of apparently transparent opacity. *Semantics & Pragmatics*, 7(7), 1–67. <http://doi.org/10.3765/sp.7.7>

- Maier, E. (2015a). Quotation and unquotation in free indirect discourse. *Mind & Language*, 30(3), 345–373. <http://doi.org/10.1111/mila.12083>
- Maier, E. (2015b). Reported speech in the transition from orality to literacy. *Glotta*, 91(1), 1–16.
- Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughn, B. E., & Flavell, J. H. (1974). The early development of inferences about the visual percepts of others. *Child Development*, 45(2), 357–366. <http://doi.org/10.2307/1127956>
- McCloskey, J. (2006). Questions and questioning in a local English. In R. Zanuttini, H. Campos, E. Herburger, & P. H. Portner (Eds.), *Cross-Linguistic Research in Syntax and Semantics: Negation, Tense and Clausal Architecture* (pp. 87–126). Washington, D.C.: Georgetown University Press.
- Meer, J. van der (2013). *Interpretation of singular personal pronouns in direct and indirect speech: A comparison between Dutch and Dutch-Frisian speakers*. Bachelor thesis, University of Groningen.
- Merleau-Ponty, M. (1964). The child's relations with others, translated by William Cobb. In J. Wild & J. M. Edie (Eds.), *The Primacy of Perception and Other Essays on Phenomenological Psychology, the Philosophy of Art, History and Politics* (pp. 96–155). Evanston: Northwestern University Press.
- Mills, A. E. (1986). *The Acquisition of Gender. A Study of English and German*. Berlin, Heidelberg: Springer.
- Mizuno, A., Liu, Y., Williams, D. L., Keller, T. a., Minshew, N. J., & Just, M. A. (2011). The neural basis of deictic shifting in linguistic perspective-taking in high-functioning autism. *Brain: A Journal of Neurology*, 134, 2422–2435. <http://doi.org/10.1093/brain/awr151>
- Moll, H., & Meltzoff, A. N. (2011). Perspective-taking and its foundation in joint attention. In J. Roessler, H. Lerman, & N. Eilan (Eds.), *Perception, Causation, and Objectivity* (pp. 286–304). Oxford: Oxford University Press.
- Moore, M. E. (2001). Third person pronoun errors by children with and without language impairment. *Journal of Communication Disorders*, 34(3), 207–228. [http://doi.org/10.1016/S0021-9924\(00\)00050-2](http://doi.org/10.1016/S0021-9924(00)00050-2)
- Mulder, H. de (2011). *Putting the pieces together: The development of theory of mind and (mental) language*. Doctoral dissertation, Utrecht University.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Drucker, K. T. (2012). *PIRLS 2011. International results in reading*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Murphy, S. (1986). Children's comprehension of deictic categories in oral and written language. *Reading Research Quarterly*, 21(2), 118–131. <http://doi.org/10.2307/747840>
- Nelson, K. (1975). The nominal shift in semantic-syntactic development. *Cognitive Psychology*, 7(4), 461–479. [http://doi.org/10.1016/0010-0285\(75\)90018-3](http://doi.org/10.1016/0010-0285(75)90018-3)
- Nordqvist, Å. (1998). Projecting speech to protagonists in oral and written narratives: A developmental study. *Psychology of Language and Communication*, 2, 37–46.
- Nordqvist, Å. (2000). The development of reported speech in speech and writing. In M.-F. Crété & E. Esperet (Eds.), *Proceedings of the 1998 European Writing Conference. Université de Poitiers, MSHS Poitiers* (pp. 333–338).
- Nordqvist, Å. (2001a). *Speech about Speech. A Developmental Study on Form and Function of Direct and Indirect Speech*. Göteborg: Kompentiet.
- Nordqvist, Å. (2001b). The use of direct and indirect speech by 1 ½- to 4-year-olds. *Psychology of Language and Communication*, 5(1), 57–66.

- Oliveira, M., & Cunha, D. A. C. (2004). Prosody as marker of direct reported speech boundary. *Proceedings of Speech Prosody 2004*, 263–266.
- Oshima, D. Y. (2006). *Perspectives in Reported Discourse*. Ph.D. thesis, Stanford University.
- Oshima-Takane, Y. (1988). Children learn from speech not addressed to them: The case of personal pronouns. *Journal of Child Language*, 15(1), 95–108. <http://doi.org/10.1017/S0305000900012071>
- Oshima-Takane, Y. (2009). Analysis of pronominal errors: A case-study. *Journal of Child Language*, 19(1), 111–131. <http://doi.org/10.1017/S0305000900013659>
- Oshima-Takane, Y., & Benaroya, S. (1989). An alternative view of pronominal errors in autistic children. *Journal of Autism and Developmental Disorders*, 19(1), 73–85. <http://doi.org/10.1007/BF02212719>
- Oshima-Takane, Y., Goodz, E., & Derevensky, J. L. (1996). Birth order effects on early language development: Do secondborn children learn from overheard speech? *Child Development*, 67(2), 621–634. <http://doi.org/10.2307/1131836>
- Oshima-Takane, Y., Takane, Y., & Shultz, T. R. (1999). The learning of first and second person pronouns in English: Network models and analysis. *Journal of Child Language*, 26(3), 545–575. <http://doi.org/10.1017/S0305000999003906>
- Özyürek, A. (1996). How children talk about a conversation. *Journal of Child Language*, 23(3), 693–714. <http://doi.org/10.1017/S0305000900009004>
- Partee, B. H. (1973). The syntax and semantics of quotation. In S. R. Anderson & P. Kiparsky (Eds.), *A Festschrift for Morris Halle* (pp. 410–418). New York: Holt, Rinehart and Winston.
- Pascual, E. (2014). *Fictive Interaction: The Conversation Frame in Thought, Language, and Discourse*. Amsterdam, Philadelphia: John Benjamins.
- Perner, J., Brandl, J., & Garnham, A. (2003). What is a perspective problem? Developmental issues in understanding belief and dual identity. *Facta Philosophica*, 5, 355–378.
- 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. <http://doi.org/10.1111/1467-8624.t01-1-00529>
- 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. [http://doi.org/10.1016/s0885-2014\(02\)00127-2](http://doi.org/10.1016/s0885-2014(02)00127-2)
- Perner, J., & Wimmer, H. (1985). “John thinks that Mary thinks that ...” Attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, 39, 437–471. [http://doi.org/10.1016/0022-0965\(85\)90051-7](http://doi.org/10.1016/0022-0965(85)90051-7)
- Perry, J. (1979). The problem of the essential indexical. *Noûs*, 13(1), 3–21. <http://doi.org/10.2307/2214792>
- Petitto, L. A. (1987). On the autonomy of language and gesture: Evidence from the acquisition of personal pronouns in American Sign Language. *Cognition*, 27(1), 1–52. [http://doi.org/10.1016/0010-0277\(87\)90034-5](http://doi.org/10.1016/0010-0277(87)90034-5)
- Philips, S. U. (1986). Reported speech as evidence in an American trial. In D. Tannen & J. E. Alatis (Eds.), *GURT '85. Languages and Linguistics: The Interdependence of Theory, Data, and Application* (pp. 154–170). Washington, D.C.: Georgetown University Press.
- Pierce, A. (1992). *Language Acquisition and Syntactic Theory: A Comparative Analysis of French and English Child Grammars*. Dordrecht: Kluwer.

- Plank, F. (1986). Über den Personenwechsel und den anderer deiktischer Kategorien in der wiedergegebenen Rede. *Zeitschrift Für Germanistische Linguistik*, 14(3), 294–308.
- Pool, C. van der. (2012). Reading picturebooks as literature: Four-to-six-year-old children and the development of literary competence. In M. Mackey (Ed.), *Picturebooks and Literary Understanding, in Honour of Lawrence Sipe* (pp. 93–106). Dordrecht: Springer.
- Potts, C. (2007). The expressive dimension. *Theoretical Linguistics*, 33(2), 165–198. <http://doi.org/10.1515/TL.2007.011>
- Quer, J. (2005). Context shift and indexical variables in sign languages. In E. Georgala & J. Howell (Eds.), *Proceedings of SALT XV* (pp. 152–168). Ithaca, NY: Cornell University.
- Quer, J. (2011). Reporting and quoting in signed discourse. In E. Brendel, J. Meibauer, & M. Steinbach (Eds.), *Understanding Quotation* (pp. 277–302). Berlin, New York: De Gruyter Mouton.
- Recanati, F. (2000). *Oratio Obliqua, Oratio Recta. An Essay on Metarepresentation*. Cambridge, MA, London, England: The MIT Press.
- Recanati, F. (2001). Open quotation. *Mind*, 110(439), 637–687. <http://doi.org/10.1093/mind/110.439.637>
- Redeker, G. (1991). Quotation in discourse. In R. van der Hout & E. Huls (Eds.), *Artikelen van de eerste Socio-linguïstische Conferentie* (pp. 341–355). Delft: Eburon.
- Reis, M. (1997). Zum syntaktischen Status unselbständiger Verbzweit-Sätze. In C. Dürscheid, K. H. Ramers, & M. Schwarz (Eds.), *Sprache im Fokus. Festschrift für Heinz Vater zum 65. Geburtstag* (pp. 121–144). Tübingen: Max Niemeyer Verlag.
- Ricard, M., Girouard, P. C., & Décarie, T. G. (1999). Personal pronouns and perspective taking in toddlers. *Journal of Child Language*, 26(3), 681–97. <http://doi.org/10.1017/S0305000999003943>
- Roberts, R. J., & Pennington, B. F. (1996). An interactive framework for examining prefrontal cognitive processes. *Developmental Neuropsychology*, 12(1), 105–126. <http://doi.org/10.1080/87565649609540642>
- Robinson, J. P., & Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. *American Educational Research Journal*, 48(2), 268–302. <http://doi.org/10.3102/0002831210372249>
- Roeck, M. de (1994). A functional typology of speech reports. In E. Engberg-Pedersen, L. F. Jakobsen, & L. S. Rasmussen (Eds.), *Function and Expression in Functional Grammar* (pp. 331–351). Berlin, New York: Mouton de Gruyter.
- Roncador, M. von (1988). *Zwischen direkter und indirekter Rede. Nichtwörtliche direkte Rede, erlebte Rede, logophorische Konstruktionen und Verwandtes*. Tübingen: Max Niemeyer Verlag.
- Rothweiler, M. (1993). *Der Erwerb von Nebensätzen im Deutschen*. Tübingen: Niemeyer.
- Ruffman, T., Slade, L., Rowlandson, K., Rumsey, C., & Garnham, A. (2003). How language relates to belief, desire, and emotion understanding. *Cognitive Development*, 18, 139–158. [http://doi.org/10.1016/S0885-2014\(03\)00002-9](http://doi.org/10.1016/S0885-2014(03)00002-9)
- Sakita, T. I. (2002). *Reporting Discourse, Tense, and Cognition*. Amsterdam: Elsevier.
- Salazar Orvig, A., Marcos, H., Morgenstern, A., Hassan, R., Leber-Marín, J., & Parès, J. (2010). Dialogical beginnings of anaphora: The use of third person pronouns before the age of 3. *Journal of Pragmatics*, 42(7), 1842–1865. <http://doi.org/10.1016/j.pragma.2009.09.020>

- Sams, J. (2010). Quoting the unspoken: An analysis of quotations in spoken discourse. *Journal of Pragmatics*, 42(11), 3147–3160. <http://doi.org/10.1016/j.pragma.2010.04.024>
- Sato, M., & Bergen, B. K. (2013). The case of the missing pronouns: Does mentally simulated perspective play a functional role in the comprehension of person? *Cognition*, 127(3), 361–74. <http://doi.org/10.1016/j.cognition.2013.02.004>
- Saxe, R. (2009). Theory of Mind (Neural basis). In W. P. Banks (Ed.), *Encyclopedia of Consciousness* (pp. 401–409). Oxford: Elsevier.
- Schlenker, P. (2003). A plea for monsters. *Linguistics and Philosophy*, 26(1), 29–120. <http://doi.org/10.1023/A:1022225203544>
- Schlenker, P. (2004). Context of thought and context of utterance: A note on free indirect discourse and the historical present. *Mind & Language*, 19(3), 279–304. <http://doi.org/10.1111/j.1468-0017.2004.00259.x>
- Schlenker, P. (2011). Indexicality and de se reports. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning* (Vol. 2, pp. 1561–1604). Berlin: Mouton de Gruyter.
- Serratrice, L., Hesketh, a., & Ashworth, R. (2015). The use of reported speech in children's narratives: A priming study. *First Language*, 35(1), 68–87. <http://doi.org/10.1177/0142723715569552>
- Sharvit, Y. (2008). The puzzle of free indirect discourse. *Linguistics and Philosophy*, 31(3), 353–395. <http://doi.org/10.1007/s10988-008-9039-9>
- Siewierska, A. (2004). *Person*. Cambridge: Cambridge University Press.
- Smiley, P., & Johnson, R. (2006). Self-referring terms, event transitivity and development of self. *Cognitive Development*, 21(3), 266–284. <http://doi.org/10.1016/j.cogdev.2006.01.008>
- Spencer, J., Smits, E.-J., & Hendriks, P. (2009). Coherent discourse solves the pronoun interpretation problem. *Journal of Child Language*, 36(1), 23–52. <http://doi.org/10.1017/S0305000908008854>
- Stawarska, B. (2009). *You and I: Dialogical Phenomenology*. Athens, Ohio: Ohio University Press.
- Stites, M. C., Luke, S. G., & Christianson, K. (2013). The psychologist said quickly, “dialogue descriptions modulate reading speed!” *Memory & Cognition*, 41(1), 137–151. <http://doi.org/10.3758/s13421-012-0248-7>
- Stukenbrock, A. (2015). *Deixis in der face-to-face-Interaktion*. Berlin, München, Boston: De Gruyter. <http://doi.org/10.1515/9783110307436>
- Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology*, 30(3), 395–402. <http://doi.org/10.1037/0012-1649.30.3.395>
- Tager-Flusberg, H. (1994). Dissociations in form and function in the acquisition of language by autistic children. In H. Tager-Flusberg (Ed.), *Constraints on Language Acquisition: Studies of Atypical Children* (pp. 175–194). Hillsdale, NJ: Lawrence Erlbaum Associates.
- 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. 289–318). Oxford: Oxford University Press.
- Tannen, D. (2007). *Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse* (2nd ed.). Cambridge: Cambridge University Press.
- Tanz, C. (1980). *Studies in the Acquisition of Deictic Terms*. Cambridge: Cambridge University Press.
- Taylor, M. (1999). *Imaginary Companions and the Children who Create them*. Oxford: Oxford University Press.

- Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.
- Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before Speech: The Beginning of Interpersonal Communication* (pp. 321–348). Cambridge: Cambridge University Press.
- Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, L. (1999). The kindergarten-path effect: studying on-line sentence processing in young children. *Cognition*, 73, 89–134. [http://doi.org/10.1016/S0010-0277\(99\)00032-3](http://doi.org/10.1016/S0010-0277(99)00032-3)
- Vandermaas-Peeler, M., Nelson, J., Bumpass, C., & Sassine, B. (2009). Social contexts of development: Parent-child interactions during reading and play. *Journal of Early Childhood Literacy*, 9(3), 295–317. <http://doi.org/10.1177/1468798409345112>
- Vandermaas-Peeler, M., Sassine, B., Price, C., & Brillhart, C. (2011). Mothers' and fathers' guidance behaviours during storybook reading. *Journal of Early Childhood Literacy*, 12(4), 415–442. <http://doi.org/10.1177/1468798411417381>
- Verbrugge, R. (2009). The facts matter, and so do computational models. *Journal of Philosophical Logic*, 38(6), 649–680.
- Vincent, D., & Perrin, L. (1999). On the narrative vs non-narrative functions of reported speech: A socio-pragmatic study. *Journal of Sociolinguistics*, 3(3), 291–313. <http://doi.org/10.1111/1467-9481.00080>
- Visker, O. (2015). *Narrative reported speech interpretation by older adults*. Bachelor thesis, University of Groningen.
- Voort, H. van der (2004). *A Grammar of Kwaza*. Berlin, New York: Mouton de Gruyter.
- Vries, M. de (2008). The representation of language within language: A syntactico-pragmatic typology of direct speech. *Studia Linguistica*, 62(1), 39–77. <http://doi.org/10.1111/j.1467-9582.2007.00142.x>
- Wade, E., & Clark, H. H. (1993). Reproduction and demonstration in quotations. *Journal of Memory and Language*, 32(6), 805–819. <http://doi.org/10.1006/jmla.1993.1040>
- Walker, A. G. (1993). Questioning young children in court: A linguistic case study. *Law and Human Behavior*, 17(1), 59–81. <http://doi.org/10.1007/BF01044537>
- Wechsler, S. (2010). What “you” and “I” mean to each other: Person indexicals, self-ascription, and theory of mind. *Language*, 86(2), 332–365. <http://doi.org/10.1353/lan.0.0220>
- Weissenborn, J., Roeper, T., & De Villiers, J. (1991). The acquisition of wh-movement in German and French. In T. L. Maxfield & B. Plunkett (Eds.), *Papers in the Acquisition of WH. Proceedings of the University of Massachusetts Roundtable, May 1990* (pp. 43–78). University of Massachusetts Occasional Papers.
- 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. <http://doi.org/10.1111/1467-8624.00304>
- Wells, B., Peppé, S., & Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language*, 31(4), 749–778. <http://doi.org/10.1017/S030500090400652X>
- Wierzbicka, A. (1974). The semantics of direct and indirect discourse. *Papers in Linguistics*, 7, 267–307. <http://doi.org/10.1080/08351817409370375>
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625–36.

- 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), 103–128. [http://doi.org/10.1016/0010-0277\(83\)90004-5](http://doi.org/10.1016/0010-0277(83)90004-5)
- Wolf, D. P., Rygh, J., & Altshuler, J. (1984). Agency and experience: Actions and states in play narratives. In I. Bretherton (Ed.), *Symbolic Play: The Development of Social Understanding* (pp. 195–217). Orlando et al.: Academic Press.
- Yao, B., Belin, P., & Scheepers, C. (2011). Silent reading of direct versus indirect speech activates voice-selective areas in the auditory cortex. *Journal of Cognitive Neuroscience*, 23(10), 3146–3152. http://doi.org/10.1162/jocn_a_00022
- Yao, B., Belin, P., & Scheepers, C. (2012). Brain “talks over” boring quotes: Top-down activation of voice-selective areas while listening to monotonous direct speech quotations. *NeuroImage*, 60(3), 1832–1842. <http://doi.org/10.1016/j.neuroimage.2012.01.111>
- Yao, B., & Scheepers, C. (2011). Contextual modulation of reading rate for direct versus indirect speech quotations. *Cognition*, 121(3), 447–453. <http://doi.org/10.1016/j.cognition.2011.08.007>
- Yule, G. (1993). Vera Hayden's dilemma, or the indirection in direct speech. In M. Eid & G. K. Iverson (Eds.), *Principles and Prediction: The Analysis of Natural Language: Papers in Honor of Gerald Sanders* (pp. 233–242). Amsterdam: John Benjamins. <http://doi.org/10.1075/cilt.98.20yul>
- Zimmermann, T. E. (1991). Kontextabhängigkeit. In H. Steger & H. E. Wiegand (Eds.), *Handbücher zur Sprach- und Kommunikationswissenschaft* (Band 6, pp. 156–229). Berlin, New York: Walter de Gruyter.
- Zwaan, R. A. (1999). Situation models: The mental leap into imagined worlds. *Current Directions in Psychological Science*, 8(1), 15–18. <http://doi.org/10.1111/1467-8721.00004>
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2), 162–185. <http://doi.org/10.1037/0033-2909.123.2.162>
- Zwart, J.-W. (1997). *Morphosyntax of Verb Movement: A Minimalist Approach to the Syntax of Dutch*. Dordrecht: Kluwer.

SAMENVATTING

INLEIDING

We kunnen over allerlei verschillende zaken praten: bijvoorbeeld het weer, de televisieserie *Game of Thrones*, of morele vragen. Maar taal maakt het ook mogelijk om over spraak zelf te praten. Het Nederlands heeft, net als veel andere talen, twee verschillende mogelijkheden om de uitingen van iemand anders weer te geven: de directe rede en de indirecte rede. Stelt u zich bijvoorbeeld eens voor dat een moeder tegen haar zoontje Peter zegt dat hij zijn kamer moet opruimen. Als Peter aan een vriendje wil vertellen wat zijn moeder tegen hem zei, dan kan hij ervoor kiezen om dat te doen in de directe rede, zoals in (1), of in de indirecte rede, zoals in (2):

- (1) Directe rede: Mama heeft gezegd: “Je moet je kamer opruimen!”
- (2) Indirecte rede: Mama heeft gezegd dat ik mijn kamer moet opruimen.

Het fundamentele verschil tussen de directe en de indirecte rede is het perspectief van waaruit iemands eerdere uitspraak wordt weergegeven. In de directe rede geeft Peter de woorden die zijn moeder eerder sprak weer vanuit háár perspectief. Dit is waarom het persoonlijk voornaamwoord tweede persoon enkelvoud *je* in (1) naar Peter verwijst, de persoon die oorspronkelijk werd aangesproken. Interpretatie van de directe rede vereist dan ook dat het perspectief van de huidige spreker (Peter) wordt verruild voor dat van de oorspronkelijke spreker (mama). Taalkundig gezien uit zich dit in een verandering in de interpretatie van verwijzende woorden, zoals persoonlijke voornaamwoorden. In de indirecte rede daarentegen geeft Peter de uitspraak die hij parafraseert weer vanuit zijn eigen perspectief en gebruikt hij daarbij het persoonlijk voornaamwoord eerste persoon enkelvoud *ik* om naar zichzelf te verwijzen.

Het onderscheid tussen de directe en de indirecte rede is echter niet zo duidelijk als het op het eerste gezicht lijkt. Taalvergelijkend onderzoek laat zien dat veel talen constructies hebben om andermans uitspraken weer te geven die niet goed zijn in te passen in de tweedeling van directe en indirecte rede (Evans, 2012). Bovendien tonen ontwikkelingsstudies aan dat kinderen tot een relatief late leeftijd moeite hebben om de directe en de indirecte rede goed van elkaar te onderscheiden (Goodell & Sachs, 1992; Hickmann, 1993; Hollebrandse, 2007). Het is niet duidelijk of dit ook een effect heeft op de manier waarop kinderen perspectief in directe en indirecte rede begrijpen. Een manier om dit te onderzoeken is om te kijken naar het begrip van voornaamwoorden, aangezien voornaamwoorden ofwel het perspectief van de gerapporteerde spreker weergeven (directe rede), ofwel dat van de rapporterende spreker (indirecte rede).

In deze dissertatie richt ik me op de volgende vragen: Is het gemakkelijker voor een luisteraar om persoonlijke voornaamwoorden te interpreteren in de directe rede of in de

indirecte rede? Hebben pragmatische factoren zoals de situatie waarin gecommuniceerd wordt invloed op het verwerken van redeweergave? Wanneer leren kinderen precies om persoonlijke voornaamwoorden correct te interpreteren en te produceren in de directe en de indirecte rede? Om deze vragen te beantwoorden heb ik een corpusstudie en drie psycholinguïstische experimenten uitgevoerd. In wat volgt, zal ik de voornaamste bevindingen van deze studies samenvatten.

CORPUSSTUDIE (HOOFDSTUK 3)

In de corpusstudie analyseerde ik het gebruik van directe en indirecte rede in de spontane taal van Nederlandstalige en Duitstalige kinderen tussen de 1 jaar en de 4½ jaar in de CHILDES database. Hieruit maakte ik op dat kinderen met het produceren van redeweergave beginnen op een leeftijd van ongeveer 2 jaar. Aanvankelijk gebruiken ze eenvoudige directe rede, waarbij ze losse woorden of geluiden aanhalen (bijvoorbeeld *Hond zegt: “waf”*). In het derde levensjaar halen kinderen ingewikkeldere uitspraken aan en beginnen ze ook de indirecte rede te gebruiken. Redeweergaven waarin kenmerken van de directe en de indirecte rede gemengd worden, zijn zeldzaam in het taalgebruik van kinderen. Zowel Nederlandstalige kinderen als Duitstalige kinderen maken aanzienlijk meer gebruik van de directe dan van de indirecte rede. Dit zou het gevolg kunnen zijn van de speelse aard van de meeste interacties die kinderen hebben. Een nadere bestudering van de discours-contexten van de redeweergave van Duitse kinderen laat inderdaad zien dat wanneer er sprake is van ‘doen alsof’ of een verhalende context, er een voorkeur is voor het gebruik van de directe rede. In contexten die meer realiteits-georiënteerd zijn, daarentegen, is er een voorkeur voor het gebruik van de indirecte rede. Verder hebben kinderen een voorkeur voor het gebruik van de indirecte rede wanneer ze naar een autoriteit verwijzen (bijvoorbeeld *Mama heeft gezegd dat je me niet mag slaan*).

EXPERIMENT 1 (HOOFDSTUK 4)

In experiment 1 behandelden we de vraag of persoonlijke voornaamwoorden gemakkelijker te interpreteren zijn in de directe rede of in de indirecte rede. We formuleerden twee mogelijke hypothesen. Aan de ene kant zouden persoonlijke voornaamwoorden moeilijker te interpreteren kunnen zijn in de directe dan in de indirecte rede omdat de directe rede een verschuiving van het perspectief van de huidige spreker naar het perspectief van de oorspronkelijke spreker vereist. Aan de andere kant zou het ook zo kunnen zijn dat kinderen niet alleen vaker de directe rede produceren, maar het ook gemakkelijker vinden om de directe rede te interpreteren. In een referent-selectie-taak keken kinderen tussen de vier en twaalf jaar en volwassenen naar animatiefilmpjes op een tablet waarin dierlijke hoofdpersonen de persoonlijke voornaamwoorden *ik*, *jij* en *hij* gebruikten in de directe rede (bijvoorbeeld *Olifant zei: “Ik krijg de voetbal”*) en in de indirecte rede (bijvoorbeeld *Olifant zei dat hij de*

voetbal krijgt). De resultaten laten zien dat zowel kinderen als volwassenen meer fouten maken en langer de tijd nemen om een keuze te maken voor een referent bij het interpreteren van persoonlijke voornaamwoorden in de directe rede in vergelijking met de indirecte rede. Hierbij is sprake van een systematisch patroon in het maken van fouten: Kinderen interpreterden persoonlijke voornaamwoorden in de directe rede vaak op dezelfde manier als in de indirecte rede. Dit duidt erop dat ze er niet in slagen om het perspectief van de huidige spreker te verruilen voor dat van de oorspronkelijke spreker. Verrassend genoeg worstelden zelfs twaalfjarigen er nog mee om persoonlijke voornaamwoorden in de directe rede correct te interpreteren. De reden hiervoor zou kunnen zijn dat bij kinderen ofwel de benodigde cognitieve vaardigheden ontbreken om van perspectief te veranderen, ofwel dat ze niet in staat zijn deze vaardigheden serieel toe te passen tijdens een interactie. De grote hoeveelheid fouten in de directe rede zou ook gerelateerd kunnen zijn aan het feit dat in dit type interactie de context waarin wordt gerapporteerd meer in het oog springt dan de context waarover wordt gerapporteerd. Anders dan volwassenen hanteren kinderen wellicht de strategie dat ze persoonlijke voornaamwoorden interpreteren met betrekking tot de meest prominente context, waarbij ze de ondubbelzinnige syntactische, semantische en prosodische signalen die de directe rede van de indirecte rede onderscheiden negeren.

EXPERIMENT 2 (HOOFDSTUK 5)

In experiment 2 onderzochten we de oorzaken van de slechte prestaties van kinderen in de directe rede in experiment 1. We gebruikten daarom soortgelijke testzinnen als in experiment 1, maar met twee belangrijke aanpassingen. Ten eerste werden de kinderen in dit experiment rechtstreeks betrokken bij de interactie als toegesprokenen. Ten tweede was de situatie in dit experiment realistischer, aangezien we gebruik maakten van acties met handpoppen in plaats van animaties op een tablet. Onze voorspelling was dat deze aanpassingen de interactie minder complex zouden maken en dat kinderen de directe rede daarom beter zouden begrijpen. Aan dit experiment namen zowel vijf- en negenjarige kinderen als volwassenen deel. We ontdekten dat in deze aangepaste situatie kinderen significant beter presteerden in het begrijpen van de directe rede dan hun leeftijdsgenoten in experiment 1. Echter, dit ging alleen op voor de persoonlijke voornaamwoorden *ik* en *hij*, maar niet voor *jij*. Als kinderen de toegesprokene zijn, hebben ze de neiging om *jij* te interpreteren als een verwijzing naar henzelf, zelfs als *jij* is ingebed in de directe rede. Dit komt overeen met Wechsler's (2010) *de se*-theorie over persoonlijke voornaamwoorden, die voorspelt dat *jij* automatisch wordt begrepen als betrekking hebbend op de luisteraar.

Daarnaast bevatte experiment 2 ook een onderdeel waarin uitdrukkingen moesten worden geproduceerd. De deelnemers moesten een uitspraak zoals *Jij krijgt de voetbal* doorgeven van de ene handpop aan de andere. Hoewel kinderen hierbij zowel directe als indirecte rede gebruikten, gaven bijna alle volwassenen de voorkeur aan de indirecte rede. In combinatie met de resultaten van de studie naar de interpretatie van directe en indirecte

rede suggereert dit dat in een discours-context waarin de uitwisseling van informatie centraal staat, de indirecte rede de voorkeur geniet. Dit geldt zowel voor het begrijpen als voor het produceren van de weergave van andermans spraak.

EXPERIMENT 3 (HOOFDSTUK 6)

In experiment 3 onderzochten we het begrijpen van redeweergave in een verhalende context. Eerdere studies hebben aangetoond dat kinderen en volwassenen in verhalen een voorkeur hebben voor het gebruik van de levendigere directe rede (Baker & Freebody, 1989; Vincent & Perrin, 1999). Omdat in een verhaal de nadruk normaal gesproken op de karakters ligt, en niet op de verteller, springt het perspectief van de karakters het meest in het oog. Een verschuiving naar het perspectief van de karakters, zoals de directe rede vereist, zou daarom gemakkelijker moeten zijn in een verhalende context dan in een context waarin informatie wordt uitgewisseld. In experiment 3 bootsten we daarom het voorlezen van een plaatjesboek na op een tablet. Vier- en zesjarige kinderen en volwassenen luisterden naar een verhaaltje met bijbehorende plaatjes. Ze moesten daarbij persoonlijke voornaamwoorden in de directe en de indirecte rede interpreteren. De resultaten laten zien dat de meeste zesjarigen in een verhalende context persoonlijke voornaamwoorden in de directe rede correct kunnen interpreteren. Zoals voorspeld, was het in deze situatie niet moeilijker om persoonlijke voornaamwoorden te interpreteren in de directe rede dan in de indirecte rede.

CONCLUSIE

Dit onderzoek had betrekking op het produceren en begrijpen van voornaamwoorden in directe en indirecte rede in taal. Ik onderzocht hoe kinderen en volwassenen dit doen aan de hand van een corpusstudie en drie psycholinguïstische experimenten.

De corpusstudie naar het spontane gebruik van directe en indirecte rede leverde bewijs dat Nederlandse en Duitse kinderen vanaf hun derde jaar in staat zijn de directe en de indirecte rede te gebruiken, en dat ze zelden kenmerken van de directe en indirecte rede door elkaar gebruiken. Deze resultaten over het produceren van redeweergave contrasteren met de bevindingen van de experimenten over het begrijpen van redeweergave. In experiment 1 bijvoorbeeld interpreteerden zelfs elfjarige persoonlijke voornaamwoorden in de directe rede in meer dan de helft van de gevallen op dezelfde manier als in de indirecte rede. Dit suggereert dat er een asymmetrie is in de verwerving van het produceren van redeweergave en het begrijpen ervan, en dat kinderen eerst de correcte productie van de weergave van andermans spraak leren, en pas daarna het correcte begrip ervan (cf. Hendriks, 2014).

In dit onderzoek heb ik drie factoren geïdentificeerd die van invloed zijn op het begrip door kinderen van persoonlijke voornaamwoorden in de directe en de indirecte rede: (i) de discours-context, (ii) de rol in de interactie en (iii) de aanwezigheid van een talig antecedent.

De eerste factor verklaart waarom kinderen, afhankelijk van de discours-context (bijvoorbeeld een context waarin informatie wordt uitgewisseld, of een verhalende context), meer of minder accuraat zijn in hun interpretatie van persoonlijke voornaamwoorden in de directe en de indirecte rede. Ik beargumenteer dat in verschillende situaties de huidige en de oorspronkelijke context meer of minder prominent zijn. Als de oorspronkelijke context bijzonder prominent is, bevordert dit de interpretatie van persoonlijke voornaamwoorden in de directe rede. Wanneer echter de huidige context bijzonder prominent is, vergemakkelijkt dit de interpretatie van persoonlijke voornaamwoorden in de indirecte rede.

De tweede factor verklaart waarom de rol die kinderen aannemen in de interactie beïnvloedt hoe ze redeweergave verwerken. Kinderen vinden het gemakkelijker om persoonlijke voornaamwoorden in de directe rede te interpreteren wanneer ze actief worden aangesproken, dan wanneer ze slechts niet-deelnemende toeschouwer zijn.

De derde factor heeft betrekking op de observatie dat een persoonlijk voornaamwoord gemakkelijker te interpreteren is als de talige context de referent expliciet noemt en er daarmee meer de aandacht op vestigt. Dit verklaart bijvoorbeeld waarom het persoonlijke voornaamwoord tweede persoon enkelvoud *jij* gemakkelijker te interpreteren is in de zin *Anita zei tegen Oscar: "Jij krijgt de voetbal"* dan in *Anita zei: "Jij krijgt de voetbal"*.

De verschillende studies in deze dissertatie hebben voor het eerst op experimentele wijze de interpretatie van persoonlijke voornaamwoorden door kinderen en volwassenen in directe versus indirecte rede onderzocht. Meer onderzoek is nodig om in kaart te brengen welke cognitieve processen er ten grondslag liggen aan het produceren en het begrijpen van de weergave van andermans uitingen. Toekomstige studies zouden zich daarnaast kunnen richten op de interpretatie van redeweergave in personen met autisme, personen met afasie en tweetaligen. Deze groepen taalgebruikers hebben wellicht om verschillende redenen moeite met het interpreteren van voornaamwoorden in redeweergave. Een verdere mogelijkheid is om de aandacht te richten op andere typen van perspectief-afhankelijke uitdrukkingen, zoals temporele (bijvoorbeeld *morgen*) of ruimtelijke (bijvoorbeeld *hier*) verwijzingen.

ABOUT THE AUTHOR

Franziska Köder was born in Schwäbisch Hall, Germany, on November 3, 1985. She studied German Philology and Philosophy at the University of Heidelberg, and, as an exchange student, at Queen Mary University of London. In January 2011, she obtained her master's degree. Subsequently, she did an internship at the Max Planck Institute for Evolutionary Anthropology in Leipzig, worked at the Institute for the German Language in Mannheim, and received a scholarship for a research stay at King's College Cambridge. In September 2011, Franziska joined the ERC project BLENDS of Emar Maier at the University of Groningen. As a PhD student, she conducted psycholinguistic experiments on children's acquisition of direct and indirect speech. Since January 2016, she is a postdoctoral researcher in the project *Acquiring figurative meanings* at the University of Oslo and investigates children's acquisition of irony.

GRONINGEN DISSERTATIONS IN LINGUISTICS (GRODIL)

1. Henriëtte de Swart (1991). *Adverbs of Quantification: A Generalized Quantifier Approach*.
2. Eric Hoekstra (1991). *Licensing Conditions on Phrase Structure*.
3. Dicky Gilbers (1992). *Phonological Networks. A Theory of Segment Representation*.
4. Helen de Hoop (1992). *Case Configuration and Noun Phrase Interpretation*.
5. Gosse Bouma (1993). *Nonmonotonicity and Categorical Unification Grammar*.
6. Peter I. Blok (1993). *The Interpretation of Focus*.
7. Roelien Bastiaanse (1993). *Studies in Aphasia*.
8. Bert Bos (1993). *Rapid User Interface Development with the Script Language Gist*.
9. Wim Kosmeijer (1993). *Barriers and Licensing*.
10. Jan-Wouter Zwart (1993). *Dutch Syntax: A Minimalist Approach*.
11. Mark Kas (1993). *Essays on Boolean Functions and Negative Polarity*.
12. Ton van der Wouden (1994). *Negative Contexts*.
13. Joop Houtman (1994). *Coordination and Constituency: A Study in Categorical Grammar*.
14. Petra Hendriks (1995). *Comparatives and Categorical Grammar*.
15. Maarten de Wind (1995). *Inversion in French*.
16. Jelly Julia de Jong (1996). *The Case of Bound Pronouns in Peripheral Romance*.
17. Sjoukje van der Wal (1996). *Negative Polarity Items and Negation: Tandem Acquisition*.
18. Anastasia Giannakidou (1997). *The Landscape of Polarity Items*.
19. Karen Lattewitz (1997). *Adjacency in Dutch and German*.
20. Edith Kaan (1997). *Processing Subject-Object Ambiguities in Dutch*.
21. Henny Klein (1997). *Adverbs of Degree in Dutch*.
22. Leonie Bosveld-de Smet (1998). *On Mass and Plural Quantification: The case of French 'des'/'du'-NPs*.
23. Rita Landeweerd (1998). *Discourse semantics of perspective and temporal structure*.
24. Mettina Veenstra (1998). *Formalizing the Minimalist Program*.
25. Roel Jonkers (1998). *Comprehension and Production of Verbs in aphasic Speakers*.
26. Erik F. Tjong Kim Sang (1998). *Machine Learning of Phonotactics*.
27. Paulien Rijkhoek (1998). *On Degree Phrases and Result Clauses*.
28. Jan de Jong (1999). *Specific Language Impairment in Dutch: Inflectional Morphology and Argument Structure*.
29. H. Wee (1999). *Definite Focus*.
30. Eun-Hee Lee (2000). *Dynamic and Stative Information in Temporal Reasoning: Korean tense and aspect in discourse*.
31. Ivelin P. Stoianov (2001). *Connectionist Lexical Processing*.
32. Klarien van der Linde (2001). *Sonority substitutions*.
33. Monique Lamers (2001). *Sentence processing: using syntactic, semantic, and thematic information*.
34. Shalom Zuckerman (2001). *The Acquisition of "Optional" Movement*.
35. Rob Koeling (2001). *Dialogue-Based Disambiguation: Using Dialogue Status to Improve Speech Understanding*.
36. Esther Ruigendijk (2002). *Case assignment in Agrammatism: a cross-linguistic study*.

37. Tony Mullen (2002). *An Investigation into Compositional Features and Feature Merging for Maximum Entropy-Based Parse Selection.*
38. Nanette Bienfait (2002). *Grammatica-onderwijs aan allochtone jongeren.*
39. Dirk-Bart den Ouden (2002). *Phonology in Aphasia: Syllables and segments in level-specific deficits.*
40. Rienk Withaar (2002). *The Role of the Phonological Loop in Sentence Comprehension.*
41. Kim Sauter (2002). *Transfer and Access to Universal Grammar in Adult Second Language Acquisition.*
42. Laura Sabourin (2003). *Grammatical Gender and Second Language Processing: An ERP Study.*
43. Hein van Schie (2003). *Visual Semantics.*
44. Lilia Schürcks-Grozeva (2003). *Binding and Bulgarian.*
45. Stasinou Konstantopoulos (2003). *Using ILP to Learn Local Linguistic Structures.*
46. Wilbert Heeringa (2004). *Measuring Dialect Pronunciation Differences using Levenshtein Distance.*
47. Wouter Jansen (2004). *Laryngeal Contrast and Phonetic Voicing: A Laboratory Phonology.*
48. Judith Rispens (2004). *Syntactic and phonological processing in developmental dyslexia.*
49. Daniëlle Bougairé (2004). *L'approche communicative des campagnes de sensibilisation en santé publique au Burkina Faso: Les cas de la planification familiale, du sida et de l'excision.*
50. Tanja Gaustad (2004). *Linguistic Knowledge and Word Sense Disambiguation.*
51. Susanne Schoof (2004). *An HPSG Account of Nonfinite Verbal Complements in Latin.*
52. M. Begoña Villada Moirón (2005). *Data-driven identification of fixed expressions and their modifiability.*
53. Robbert Prins (2005). *Finite-State Pre-Processing for Natural Language Analysis.*
54. Leonoor van der Beek (2005). *Topics in Corpus-Based Dutch Syntax.*
55. Keiko Yoshioka (2005). *Linguistic and gestural introduction and tracking of referents in L1 and L2 discourse.*
56. Sible Andringa (2005). *Form-focused instruction and the development of second language proficiency.*
57. Joanneke Prenger (2005). *Taal telt! Een onderzoek naar de rol van taalvaardigheid en tekstbegrip in het realistisch wiskundeonderwijs.*
58. Neslihan Kansu-Yetkiner (2006). *Blood, Shame and Fear: Self-Presentation Strategies of Turkish Women's Talk about their Health and Sexuality.*
59. Mónika Z. Zempléni (2006). *Functional imaging of the hemispheric contribution to language processing.*
60. Maartje Schreuder (2006). *Prosodic Processes in Language and Music.*
61. Hidetoshi Shiraishi (2006). *Topics in Nivkh Phonology.*
62. Tamás Biró (2006). *Finding the Right Words: Implementing Optimality Theory with Simulated Annealing.*
63. Dieuwke de Goede (2006). *Verbs in Spoken Sentence Processing: Unraveling the Activation Pattern of the Matrix Verb.*
64. Eleonora Rossi (2007). *Clitic production in Italian agrammatism.*
65. Holger Hopp (2007). *Ultimate Attainment at the Interfaces in Second Language Acquisition: Grammar and Processing.*
66. Gerlof Bouma (2008). *Starting a Sentence in Dutch: A corpus study of subject- and object-fronting.*
67. Julia Klitsch (2008). *Open your eyes and listen carefully. Auditory and audiovisual speech perception and the McGurk effect in Dutch speakers with and without aphasia.*
68. Janneke ter Beek (2008). *Restructuring and Infinitival Complements in Dutch.*
69. Jori Mur (2008). *Off-line Answer Extraction for Question Answering.*

70. Lonneke van der Plas (2008). *Automatic Lexico-Semantic Acquisition for Question Answering*.
71. Arjen Versloot (2008). *Mechanisms of Language Change: Vowel reduction in 15th century West Frisian*.
72. Ismail Fahmi (2009). *Automatic term and Relation Extraction for Medical Question Answering System*.
73. Tuba Yarbay Duman (2009). *Turkish Agrammatic Aphasia: Word Order, Time Reference and Case*.
74. Maria Trofimova (2009). *Case Assignment by Prepositions in Russian Aphasia*.
75. Rasmus Steinkrauss (2009). *Frequency and Function in WH Question Acquisition. A Usage-Based Case Study of German L1 Acquisition*.
76. Marjolein Deunk (2009). *Discourse Practices in Preschool. Young Children's Participation in Everyday Classroom Activities*.
77. Sake Jager (2009). *Towards ICT-Integrated Language Learning: Developing an Implementation Framework in terms of Pedagogy, Technology and Environment*.
78. Francisco Dellatorre Borges (2010). *Parse Selection with Support Vector Machines*.
79. Geoffrey Andogah (2010). *Geographically Constrained Information Retrieval*.
80. Jacqueline van Kruiningen (2010). *Onderwijsontwerp als conversatie. Probleemoplossing in interprofessioneel overleg*.
81. Robert G. Shackleton (2010). *Quantitative Assessment of English-American Speech Relationships*.
82. Tim Van de Cruys (2010). *Mining for Meaning: The Extraction of Lexico-semantic Knowledge from Text*.
83. Therese Leinonen (2010). *An Acoustic Analysis of Vowel Pronunciation in Swedish Dialects*.
84. Erik-Jan Smits (2010). *Acquiring Quantification. How Children Use Semantics and Pragmatics to Constrain Meaning*.
85. Tal Caspi (2010). *A Dynamic Perspective on Second Language Development*.
86. Teodora Mehotcheva (2010). *After the fiesta is over. Foreign language attrition of Spanish in Dutch and German Erasmus Student*.
87. Xiaoyan Xu (2010). *English language attrition and retention in Chinese and Dutch university students*.
88. Jelena Prokić (2010). *Families and Resemblances*.
89. Radek Šimik (2011). *Modal existential wh-constructions*.
90. Katrien Colman (2011). *Behavioral and neuroimaging studies on language processing in Dutch speakers with Parkinson's disease*.
91. Siti Mina Tamah (2011). *A Study on Student Interaction in the Implementation of the Jigsaw Technique in Language Teaching*.
92. Aletta Kwant (2011). *Geraakt door prentenboeken. Effecten van het gebruik van prentenboeken op de sociaal-emotionele ontwikkeling van kleuters*.
93. Marlies Kluck (2011). *Sentence amalgamation*.
94. Anja Schüppert (2011). *Origin of asymmetry: Mutual intelligibility of spoken Danish and Swedish*.
95. Peter Nabende (2011). *Applying Dynamic Bayesian Networks in Transliteration Detection and Generation*.
96. Barbara Plank (2011). *Domain Adaptation for Parsing*.
97. Cagri Coltekin (2011). *Catching Words in a Stream of Speech: Computational simulations of segmenting transcribed child-directed speech*.
98. Dörte Hessler (2011). *Audiovisual Processing in Aphasic and Non-Brain-Damaged Listeners: The Whole is More than the Sum of its Parts*.

99. Herman Heringa (2012). *Positional constructions*.
100. Diana Dimitrova (2012). *Neural Correlates of Prosody and Information Structure*.
101. Harwintha Anjarningsih (2012). *Time Reference in Standard Indonesian Agrammatic Aphasia*.
102. Myrte Gosen (2012). *Tracing learning in interaction. An analysis of shared reading of picture books at kindergarten*.
103. Martijn Wieling (2012). *A Quantitative Approach to Social and Geographical Dialect Variation*.
104. Gisi Cannizzaro (2012). *Early word order and animacy*.
105. Kostadin Cholakov (2012). *Lexical Acquisition for Computational Grammars. A Unified Model*.
106. Karin Beijering (2012). *Expressions of epistemic modality in Mainland Scandinavian. A study into the lexicalization-grammaticalization-pragmaticalization interface*.
107. Veerle Baaijen (2012). *The development of understanding through writing*.
108. Jacolien van Rij (2012). *Pronoun processing: Computational, behavioral, and psychophysiological studies in children and adults*.
109. Ankelen Schippers (2012). *Variation and change in Germanic long-distance dependencies*.
110. Hanneke Loerts (2012). *Uncommon gender: Eyes and brains, native and second language learners, & grammatical gender*.
111. Marjoleine Sloos (2013). *Frequency and phonological grammar: An integrated approach. Evidence from German, Indonesian, and Japanese*.
112. Aysa Arylova (2013). *Possession in the Russian clause. Towards dynamicity in syntax*.
113. Daniël de Kok (2013). *Reversible Stochastic Attribute-Value Grammars*.
114. Gideon Kotzé (2013). *Complementary approaches to tree alignment: Combining statistical and rule-based methods*.
115. Fridah Katshemererwe (2013). *Computational Morphology and Bantu Language Learning: an Implementation for Runyakitara*.
116. Ryan C. Taylor (2013). *Tracking Referents: Markedness, World Knowledge and Pronoun Resolution*.
117. Hana Smiskova-Gustafsson (2013). *Chunks in L2 Development: A Usage-based Perspective*.
118. Milada Walková (2013). *The aspectual function of particles in phrasal verbs*.
119. Tom O. Abuom (2013). *Verb and Word Order Deficits in Swahili-English bilingual agrammatic speakers*.
120. Gülsen Yilmaz (2013). *Bilingual Language Development among the First Generation Turkish Immigrants in the Netherlands*.
121. Trevor Benjamin (2013). *Signaling Trouble: On the linguistic design of other-initiation of repair in English conversation*.
122. Nguyen Hong Thi Phuong (2013). *A Dynamic Usage-based Approach to Second Language Teaching*.
123. Harm Brouwer (2014). *The Electrophysiology of Language Comprehension: A Neurocomputational Model*.
124. Kendall Decker (2014). *Orthography Development for Creole Languages*.
125. Laura S. Bos (2015). *The Brain, Verbs, and the Past: Neurolinguistic Studies on Time Reference*.
126. Rimke Groenewold (2015). *Direct and indirect speech in aphasia: Studies of spoken discourse production and comprehension*.
127. Huiping Chan (2015). *A Dynamic Approach to the Development of Lexicon and Syntax in a Second Language*.

128. James Griffiths (2015). *On appositives*.
129. Pavel Rudnev (2015). *Dependency and discourse-configurationality: A study of Avar*.
130. Kirsten Kolstrup (2015). *Opportunities to speak. A qualitative study of a second language in use*.
131. Güliz Güneş (2015). *Deriving Prosodic structures*.
132. Cornelia Lahmann (2015). *Beyond barriers. Complexity, accuracy, and fluency in long-term L2 speakers' speech*.
133. Sri Wachyuni (2015). *Scaffolding and Cooperative Learning: Effects on Reading Comprehension and Vocabulary Knowledge in English as a Foreign Language*.
134. Albert Walsweer (2015). *Ruimte voor leren. Een etnografisch onderzoek naar het verloop van een interventie gericht op versterking van het taalgebruik in een knowledge building environment op kleine Friese basisscholen*.
135. Aleyda Lizeth Linares Calix (2015). *Raising Metacognitive Genre Awareness in L2 Academic Readers and Writers*.
136. Fathima Mufeeda Irshad (2015). *Second Language Development through the Lens of a Dynamic Usage-Based Approach*.
137. Oscar Strik (2015). *Modelling analogical change. A history of Swedish and Frisian verb inflection*.
138. He Sun (2015). *Predictors and stages of very young child EFL learners' English development in China*.
139. Marieke Haan (2015). *Mode Matters. Effects of survey modes on participation and answering behavior*.
140. Nienke Houtzager (2015). *Bilingual advantages in middle-aged and elderly populations*.
141. Noortje Joost Venhuizen (2015). *Projection in Discourse: A data-driven formal semantic analysis*.
142. Valerio Basile (2015). *From Logic to Language: Natural Language Generation from Logical Forms*.
143. Jinxing Yue (2016). *Tone-word Recognition in Mandarin Chinese: Influences of lexical-level representations*.
144. Seçkin Arslan (2016). *Neurolinguistic and Psycholinguistic Investigations on Evidentiality in Turkish*.
145. Rui Qin (2016). *Neurophysiological Studies of Reading Fluency. Towards Visual and Auditory Markers of Developmental Dyslexia*.
146. Kashmiri Stec (2016). *Visible Quotation: The Multimodal Expression of Viewpoint*.
147. Yinxing Jin (2016). *Foreign language classroom anxiety: A study of Chinese university students of Japanese and English over time*.
148. Joost Hurkmans (2016). *The Treatment of Apraxia of Speech. Speech and Music Therapy, an Innovative Joint Effort*.
149. Franziska Köder (2016). *Between direct and indirect speech: The acquisition of pronouns in reported speech*.

GRODIL

Center for Language and Cognition Groningen (CLCG)
 P.O. Box 716
 9700 AS Groningen
 The Netherlands

