



University of Groningen

Children mix direct and indirect speech

Köder, Franziska; Maier, Emar

Published in: Journal of Child Language

DOI: 10.1017/S0305000915000318

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., & Maier, E. (2016). Children mix direct and indirect speech: evidence from pronoun comprehension. *Journal of Child Language*, *43*(4), 843-866. https://doi.org/10.1017/S0305000915000318

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).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverneamendment.

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.

J. Child Lang. 43 (2016), 843–866. © Cambridge University Press 2015 doi:10.1017/S0305000915000318

Children mix direct and indirect speech: evidence from pronoun comprehension

FRANZISKA KÖDER*

Faculty of Philosophy, Center for Language and Cognition Groningen (CLCG), University of Groningen, the Netherlands

$A\,N\,D$

EMAR MAIER

Faculty of Philosophy, University of Groningen, the Netherlands

(Received 9 October 2014 – Revised 30 January 2015 – Accepted 29 May 2015 – First published online 30 June 2015)

ABSTRACT

This study investigates children's acquisition of the distinction between direct speech (*Elephant said*, "I get the football") and indirect speech (*Elephant said that he gets the football*), by measuring children's interpretation of first, second, and third person pronouns. Based on evidence from various linguistic sources, we hypothesize that the direct-indirect distinction is acquired relatively late. We also predict more mistakes for third person pronouns compared to first and second person pronouns. We tested 136 Dutch-speaking children between four and twelve in a referent selection task and found that children interpret pronouns in direct speech predominantly as in indirect speech, supporting our hypothesis about a late acquisition of the direct-indirect distinction. In addition, we found differences between I, you, and he that deviate from a simple first and second vs. third person split. We discuss our results in the light of cross-linguistic findings of direct-indirect mixing.

^[*] This research was conducted at the University of Groningen, the Netherlands. The work was supported by the EU under FP7, ERC Starting Grant 263890-BLENDS (Emar Maier). We are grateful to the children and adults who participated in our study and to the Westerschool (Wildervank), obs De Sleutel (Wildervank) and obs De Sterrensteen (Groningen) for their cooperation. We thank Jelmer van der Linde and Martijn Luinstra for technical and graphic assistance with the stimuli, Iris Strangmann for help with the data collection, and Petra Hendriks, the members of the Acquisition Lab at the University of Groningen, the editors, and two reviewers for helpful comments and suggestions. Address for correspondence: Franziska Köder, Faculty of Philosophy, Oude Boteringestraat 52, 9712 GL Groningen, The Netherlands. e-mail: f.koder@rug.nl

KÖDER AND MAIER

INTRODUCTION

Imagine that Mary comes up to me at a conference uttering the words *You're a genius!*. If I want to report this to you the next day, I can choose to do so using either a direct or an indirect speech report:

- (1) a. DIRECT SPEECH: Mary said, "You're a genius!".
 - b. INDIRECT SPEECH: Mary said that I'm a genius.

In this paper we investigate the acquisition of the distinction between direct and indirect speech. More specifically, we want to know when children acquire the ability to reliably tell apart a direct from an indirect report in Dutch. To measure this ability, we look at children's interpretation of pronouns embedded in report complements. As we see in (I), when embedded inside direct speech, the interpretation of pronouns is 'shifted': *You* in my utterance of (Ia) does not refer to my current addressee, you, but to the addressee of Mary's original utterance, me. In the interpretation of indirect speech, (Ib), no such reference shifting occurs: I simply refers to the current speaker, me.

The main hypothesis of this study is that children acquire a strict directindirect distinction relatively late. This is based on cross-linguistic and developmental evidence suggesting that the distinction between direct and indirect speech is less rigid than traditionally assumed. In our experiment, a referent selection task, we expect children to make systematic mistakes in their evaluation of pronouns in speech reports.

THEORETICAL BACKGROUND

The direct-indirect distinction

At first sight, direct and indirect speech appear to be two fundamentally distinct ways of reporting what someone said. In direct speech we report someone's words by reproducing them more or less verbatim, while in indirect speech we present what was said from our own perspective. Accordingly, in theoretical linguistics, direct and indirect speech are treated rather differently. In syntax, direct speech complements are typically analyzed as independent main clauses, while indirect speech involves clausal subordination (Banfield, 1982; De Vries, 2008). In formal semantics, direct speech is typically analyzed as a form of quotation (i.e. a form of reference to linguistic objects like sentences or utterances), while indirect speech is analyzed as an intensional operator (i.e. a quantifier over possible worlds), on a par with attitude ascriptions (*believes that*) and modal operators (*it is possible that*) (Kaplan, 1989; Maier, 2009; Schlenker, 2003; Zimmermann, 1991).

The assumed fundamental distinction between the two report types is visible at different levels of linguistic analysis. In the remainder of this

section we review the main prosodic, syntactic, and semantic characteristics supporting the split. We will illustrate with the language of our study, Dutch. The reason we chose Dutch was that, compared to English, it has extensive syntactic marking of the direct–indirect distinction, leaving very little room for ambiguity. We relied on the characteristics identified below to create clearly distinct direct and indirect speech stimuli.

Prosody/phonology. In spoken language, direct speech reports are marked by a greater overall pitch range than indirect speech reports, and are more likely to be preceded by an intonational phrase break (Jansen, Gregory & Brenier, 2001; Oliveira & Cunha, 2004). Speakers also tend to change their volume, rhythm, and speech rate in a direct quotation (Klewitz & Couper-Kuhlen, 1999). In addition to this prosodic marking, direct speech may be signaled gesturally, for instance, by airquotes or by mimicking the reported speaker's posture or physique (Clark & Gerrig, 1990). In writing, quotations are usually enclosed in quotation marks or italicized.

Syntax. In Dutch, word order differentiates between direct and indirect speech. Declarative sentences like *Ik ben blij* 'I am happy' are reported in their original verb-second word order in direct speech, but are transformed to verb-final word order in indirect speech.

(2) a. DIRECT: Anna zei, "Ik ben blij".
'Anna said, "I am happy".'
b. INDIRECT: Anna zei dat ze blij was.
'Anna said that she was happy.'

The pair in (2) reveals another useful cue for distinguishing direct and indirect speech, namely the complementizer *dat* 'that', which is obligatory in indirect reports of assertions in Dutch.

The syntax-semantics interface. A number of distinguishing characteristics of direct and indirect speech lie at the intersection of syntax and semantics. For instance, Dutch indirect speech exhibits sequence of tense, that is, the tense marking of the embedded clause is adjusted to match that of the matrix verb. We can see this in (2b), where *I am happy* (present tense) is adjusted to *that she was happy* (simple past) because the matrix clause is past tense (*said*). In addition, while direct speech blocks *wh*-extraction out of the quotation (3a), indirect speech allows extraction out of the complement (3b) (Schlenker, 2011):

(3) ORIGINAL: Jullie moeten je kamer opruimen. 'You guys have to tidy up your room.'
a. DIRECT: *Wat zei mama, "Jullie moeten"? 'What did mom say, "You guys have to"?'
b. INDIRECT: Wat zei mama dat we moeten doen? 'What did mom say (that) we have to do?'

Semantics. A purely semantic criterion for distinguishing direct and interpretation of indirect speech is the indexicals, that is, context-dependent expressions like I, you, here, and tomorrow. In both Dutch and English, indexicals receive different interpretations in direct and indirect speech, as witness the contrast between Mary said, "I'm a genius!" and Mary said that I'm a genius. In semantic terms, indexicals in direct speech get their reference from the reported speaker's utterance context, while in indirect speech they are evaluated with respect to the actual utterance context of the report. Hence, direct speech may be characterized semantically as involving a context shift (Recanati, 2000).

In this paper we use the REPORTING-REPORTED terminology to distinguish the two speech acts and utterance contexts relevant in a speech report. When John utters *Mary said*, "*I'm a genius*!", the reporting utterance is that whole utterance (*Mary said*, "*I'm a genius*!") and the reported utterance is Mary's original statement (*I'm a genius*!) that presumably occurred earlier. Consequently, the reporting context is the context where John is speaking to someone about Mary while the reported context is the earlier context where Mary is talking about herself.

Constructions between direct and indirect speech

We have seen how to distinguish direct and indirect speech in Dutch by various prosodic, syntactic, and semantic cues. This supports the traditional dichotomy in which languages provide two entirely distinct linguistic mechanisms to report what someone said; one a form of quotation and the other an intensional embedding. On closer examination, it turns out that the matter is more complicated in that there are many reporting constructions within and beyond Dutch that fall somewhere in between direct and indirect speech.

A well-known counter-example to the direct-indirect dichotomy is free indirect discourse, a form of report typically found in literary narrative (Banfield, 1982).

(4) 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.'

Syntactically, (4) looks like direct speech, as the content of the report is represented in the form of a main clause question. On the other hand, the sequence of tense and the third person pronoun *hij* 'he' point in the direction of indirect speech, as the protagonist's original utterance must have been something like *Who do you think you are!*?

A different kind of mix that also involves direct speech word order and indirect speech pronoun interpretation occurs in various Germanic languages. In Hiberno-English, for instance, apparently indirect questions may retain main clause word order:

(5) The baritone was asked what did he think of Mrs Kearney's conduct. (from James Joyce's *Dubliners*, cited by McCloskey, 2006)

Interestingly, this form of mixed reporting, which includes the so-called embedded verb-second construction, is considered grammatical in many Germanic languages, including Danish, German, and Frisian, but not in standard Dutch (cf. Zwart, 1997, and references cited therein).

If we look beyond Germanic, we find many more mixes. Consider first sign languages, where utterances are typically reported with a construction called Role Shift. In Role Shift, the reporter shifts her body and gaze away from the current addressee while reproducing the reported utterance. This is typically considered a form of direct speech, since, for instance, the first person pronoun is shifted (i.e. referring to the reported rather than the reporting signer). However, some expressions may be adjusted to the reporting environment, as in indirect speech. For instance, in Dutch Sign Language a signer might report Martine's utterance of *I think Cruijff is the best soccer player* with a non-verbatim, partially unshifted Role Shift reproduction like (6), in which <<...>> marks the scope of the non-manual Role Shift marking.

(6) MARTINE << I THINK HE_[point at Cruijff] SOCCER BEST >> 'Martine: "I think he_[point at Cruijff] is the best soccer player".'

Similar mixes are described for other sign languages, including Danish Sign Language (Engberg-Pedersen, 1995), Catalan Sign Language (Quer, 2005), and German Sign Language (Herrmann & Steinbach, 2012).

In the typological literature we find many more languages that allow apparent mixes of direct and indirect speech. Consider the following report in Slave, an Athabaskan language spoken in Canada (Rice, 1986):

(7) Yeri Margaret segha woshi néhdi What Margaret for-me you-will-make told-you 'What did Margaret tell you to make for her?'

In (7) we see a shifted first person (*for-me*) referring to the reported speaker, Margaret, as in direct speech. However, the extracted *wh*-word indicates indirect speech. Constructions like these are well documented for Amharic (Schlenker, 2003), Zazaki (Anand & Nevins, 2003), Uyghur (Shklovsky & Sudo, 2013), and Matses (Munro, Ludwig, Sauerland & Fleck, 2012), among others.

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. For instance, Maier (2009) proposes a semantic analysis of reported speech where direct and indirect speech are just the limiting cases of a more general semantic mechanism of mixed quotation. In a similar vein, Evans (2012) proposes a canonical approach to the typology of reported speech, in which traditional direct and indirect speech are merely canonical ideals spanning a continuum of potential non-canonical forms. The results of our experiment will support such a continuum hypothesis, by showing that even speakers of a language with a very rigid directindirect dichotomy in the grammar, like Dutch, start out with a more fluid style of reporting, and acquire the adult-like distinction only very late.

Direct-indirect mixes in child language

We have established that the direct-indirect distinction is not as fundamental or strict as suggested by the traditional semantic characterization (quotation vs. intensional embedding) or the initial data from standard Dutch. So far, we have encountered mixes of direct and indirect speech by exploring other languages and registers. In this paper, we explore the developmental dimension, looking for mixes in Dutch child language.

There is already some evidence in the literature that children, both in production and in comprehension, struggle to distinguish direct and indirect reports. Based on this evidence, which we review below, we will hypothesize that Dutch children acquire a clear-cut direct–indirect distinction relatively late, despite the rigid distinction typically assumed in the grammar of standard Dutch.

Several corpus studies have shown that children begin to produce both direct and indirect speech from about two to three years of age (Ely & McCabe, 1993 (English); Köder, 2013 (German and Dutch); Nordqvist, 2001 (Swedish)). In Dutch and German, children between the ages of 1;1 and 4;6 and their caregivers produce predominantly direct speech in natural interactions (Köder, 2013). In many studies, various direct-indirect mixes have been observed. This includes apparent errors applying sequence of tense, and mixes involving pronouns and other indexicals (Goodell & Sachs, 1992; Hickmann, 1993; Nordqvist, 2001). A typical example of the latter is (8):

(8) And the birdie said that I feel much prettier knowing how long my beak is. (Goodell & Sachs, 1992, p. 407)

The child (age 8;6) who produced (8) uses the complementizer *that* of indirect speech, but copies the pronouns I and my from the original dialogue, as in direct speech.

In addition to these production data, some psycholinguistic experiments have provided evidence that children also fail to properly distinguish direct and indirect speech in comprehension. First, even twelve-year-old children show a non-adultlike ability to detect prosodic cues for direct and indirect speech (Hewlett, Kelsey & Lickley, 2003). Second, children have been claimed to allow *wh*-extraction out of direct speech complements in languages such as English, German, and French (Hollebrandse, 2007; Weissenborn, Roeper & De Villiers, 1991). For instance, Hollebrandse (2007) shows that five- and six-year-old English-speaking children may interpret (9) as asking for a manner of bike riding while interpreting *I* as referring to Deanne:

(9) How did Deanne ask, "Can I ride a bike?"

In other words, children interpreted (9) via a long *wh*-movement out of the complement, which indicates indirect speech, while they interpreted the pronoun with respect to a shifted context, which indicates direct speech. This means that children's interpretation of reports shows signs of both direct and indirect speech interpretation.¹

In sum, there is some evidence from production and comprehension that children's development of a direct-indirect distinction extends throughout the childhood period. However, while direct-indirect mixings are well documented in children's production of speech reports, very little is known about how and when children acquire adultlike performance in the comprehension of speech reports. So far, only children's difficulties in obeying the *wh*-movement block in direct speech have been established. In this study, we focus on a different criterion to distinguish direct and indirect speech: the interpretation of indexicals, more specifically, deictic pronouns.

Hypotheses

We designed a simple referent selection task to explore children's interpretation of first, second, and third person singular pronouns in direct

¹ Hollebrandse (2007) 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. But how can we establish that, given that prosody is no reliable guide for children (cf. Hewlett et al., 2003), and that some dialects and registers of English allow inverted word order in indirect questions (cf. McCloskey, 2006)? According to Hollebrandse, the shifted interpretation of the pronoun establishes that children interpreted the sentence as a direct report. But one could just as easily maintain that the *wh*-movement establishes an 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. Our position sidesteps these issues. We do not cling to any one characteristic as showing that a given report is either direct or indirect. Instead, we allow the possibility of mixed reports that exhibit characteristics of both direct and indirect speech.

Downloaded from https://www.cambridge.org/core. Rijksuniversiteit Groningen, on 13 Jun 2018 at 13:59:08, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0305000915000318

and indirect speech. Based on the considerations above we put forward the following two specific hypotheses.

The main hypothesis concerns the difference between direct and indirect speech reports. We predict that children fail to clearly separate these two reporting modes, despite clear syntactic 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 that children will be especially prone to interpreting direct speech as indirect speech, because interpreting direct speech requires a cognitively demanding perspective shift from the reporting to the reported utterance context (Köder, Maier & Hendriks, 2015).

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; 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.

METHOD

Participants

The participants of this study were 136 monolingual Dutch-speaking children between ages 4;1 and 12;8 (see Table 1). 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, thirty-three adult native speakers of Dutch – mostly students – participated without compensation. The adult data are discussed in detail as part of a larger population in Köder *et al.* (2015).

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
II	11;11	11;4-12;8	19	9/10
Adults	22	18-32	33	21/12

TABLE I. Participants

All participants were tested individually in a quiet room at the school or university.

Stimuli and procedure

The experiment is designed as a child-friendly game that participants played on a tablet. It uses an Android application linked to a small off-line webpage. JavaScript controls the scenes and interactive responses, and times them using the system time in milliseconds. The experimental game is called 'Which animal gets which object?', 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, eighteen 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 2). First fifteen no report items were presented, followed by thirty 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 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

² An on-line version of the game can be played at <http://tinyurl.com/o7bburc> (Google Chrome browser recommended).

KÖDER AND MAIER

Condition	Test sentences	n
No report	Ik/Jij/Hij krijg(t) de/het [object]. 'I/You/He get(s) the [object].'	15
Speech report		30
a. Direct	Aap/Olifant/Hond zei, "Ik/Jij/Hij krijg(t) de/het [object]". 'Monkev/Elephant/Dog said, "I/You/He get(s) the [object]".'	(15)
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].'	(15)

TABLE 2. Overview of test conditions

addition, direct speech sentences have a 800 ms break between reporting clause and quotation, and include a change of pitch in the quotation.

All test sentences contain either a first, second, or third person singular pronoun (ik 'I', jij 'you', or hij 'he'). 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. In natural interactions, third person pronouns are usually used with a prior linguistic antecedent or additional extralinguistic cues such as pointing or eve-gaze (Diessel, 2012; Levinson, 2004; Salazar Orvig, Marcos, Morgenstern, Hassan, Leber-Marin & Parès, 2010). We decided against the inclusion of additional information in the case of third person pronouns in order to keep the third person stimuli uniform with the first and second person ones. Note that this could be at the expense of the naturalness of the third person stimuli in our experiment.

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 the participating 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* (Figure 1a). After the utterance, all three animals are highlighted in yellow and a basket appears in front of them (Figure 1b). The participants' task was to select the recipient of the book by touching him. In our example, the correct choice would be the dog. After selection, the object jumps into



Fig. 1. Example of a no report item. 1a. Uttering of no report sentence (e.g. *He gets the book.*). 1b. Selection phase.

the basket of the chosen animal. The software records accuracy of pronoun interpretation.

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 (Figure 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 construction (Figure 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.

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 Figure 2a). By contrast, 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 Figure 2b). 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.



Fig. 2. Example of a speech report item. 2a. Unintelligible whispering (reported context). 2b. Uttering of speech report (e.g. *Elephant said*, "*I get the football*".) (reporting context). 2c. Selection phase.

We counterbalanced the participant-roles that the animal protagonists assume in the scenes, their spatial position, and the types of sentences they utter. The experiment took approximately 10 to 15 minutes to complete.

RESULTS

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

No report condition

Figure 3 shows the percentage of correct reference assignment for the pronouns ik 'I', jij 'you', and hij 'he' in the no report condition – divided into age groups. 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.



Fig. 3. Percentage of correct pronoun interpretation in non-reportative sentences per age group. Error bars indicate 95% confidence intervals.

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 3). 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.

Speech report condition

Figures 4 and 5 show the percentage of correct pronoun interpretation in direct speech (Figure 4) and indirect speech (Figure 5).

KÖDER AND MAIER

	Estimate	SE	z value	p value
(Intercept)	9:3688	1·7674	5·301	< .001
jij 'you' vs. ik 'I'	0:1331	2·8483	0·047	.963
hij 'he' vs. ik 'I'	-22:1586	2·1274	-10·416	< .001
EXPERIENCE	0:6978	0·1736	4·019	< .001

 TABLE 3. Fixed-effect coefficients of the model fitted to children's accuracy of pronoun interpretation in no report condition

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

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, jij, 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).

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, Bretz & Westfall, 2008). The results are presented in Table 4. Children made fewer mistakes when the same pronoun occurs in indirect speech in contrast to direct speech (for *ik*: p < .001; for *jij*: 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 jij (p < 001). The mean accuracy of ikand *jij* 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 *jij* (p < 0.01) and *hij* (p < 0.01). Surprisingly, children were better in selecting the correct referent for third person hij than for second person *jij* in direct speech (p = 0.00). 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.

Direct-indirect mixes. We call it a direct-indirect mix when children evaluate a pronoun in a report with respect to the incorrect utterance context, that is,



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



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

in direct speech with respect to the reporting utterance context, and in indirect speech with respect to the reported utterance 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*. Ninety-nine percent (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 different from chance (t(138) = 1.45, p = .15). This means that children predominantly interpret pronouns in direct speech as in indirect speech, but not the other way around.

		<u>an</u>	1	· 1
Linear hypotheses	Estimate	SE	z value	<i>p</i> value
Indirect ik – Direct ik = 0	7.05	·55	12.81	< .001
Indirect jij – Direct jij = 0	8.68	·58	15.01	< .001
Indirect $hij - Direct hij = 0$	6.58	.52	12.76	< .001
Indirect ik – Indirect jij = 0	-0.30	.32	-0.63	·986
Indirect ik – Indirect hij = 0	1.24	·27	4.56	< .001
Indirect <i>jij</i> – Indirect <i>hij</i> = 0	1.44	·29	5.05	< .001
Direct ik – Direct jij = 0	1.43	·20	6.96	< .001
Direct ik – Direct $hij = 0$	0.77	·18	4.26	< .001
Direct jij – Direct hij = 0	-o·65	·21	-3.08	.010

TABLE 4. Multiple comparisons of means (Tukey contrasts) for children's accuracy of pronoun interpretation in speech report condition

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 for pronouns embedded in direct speech. This is in line with the findings of Köder *et al.* (2015), 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 adultlike 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 utterance 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, 2001). 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 lexical 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 utterance contexts. The elephant can, for instance, be the addressee in the reported utterance context and the speaker in the reporting utterance context. Children usually understand this type of perspective difference at around the age of four (Perner, Brandl & Garnham, 2003). More complex perspective-taking tasks, involving, for instance, second-order beliefs, are mastered between five and seven (Perner & Wimmer, 1985; Sullivan, Zaitchik & Tager-Flusberg, 1994). Consequently, this second explanation can account only for the younger age groups.

The third and, in our view, most important explanation is that the reporting utterance 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 various languages. As discussed in the

'Introduction', 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 direct speech correspond to how Dutch children interpret first, second, and third person pronouns embedded in Dutch direct speech. Following Evans' 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 reaction times for direct speech items (Köder et al., 2015).

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 yet at an adult level 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 do not yet have the necessary cognitive resources to correct their initial interpretation.

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 utterance 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 (Brener, 1983; Charney, 1980; Deutsch & Pechmann, 1978; Legendre & Smolensky, 2012; Murphy, 1986), we found support for Lyons' (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 made significantly more mistakes for *hij* 'he' as compared to *ik* 'I' and *jij* 'you'. 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). As noted in the 'Methods' section above, 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.

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 extralinguistically 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 anaphoric *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 extralinguistic 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 findings 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. Köder et al. (2015) report the same result for adults. They explain the apparent ease of ik in direct speech with 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 is 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.

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 don't 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 distinguish 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 as in indirect speech, while errors in the opposite direction (interpreting pronouns in indirect speech as in indirect speech as in indirect speech as in indirect speech as in speech) were rare. This indicates that in our task, an unshifted interpretation of pronouns, as in indirect speech and non-reportative sentences, is the default for children. Similarly to signers and speakers of languages such as Kwaza, Slave, and Nez Perce, Dutch children seem to

Downloaded from https://www.cambridge.org/core. Rijksuniversiteit Groningen, on 13 Jun 2018 at 13:59:08, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0305000915000318

'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, extralinguistic factors such as the salience of the reporting utterance 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 (2015) 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 adultlike direct–indirect dichotomy.

REFERENCES

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6), 716–23.
- Anand, P. (2006). De de se. Unpublished doctoral dissertation, MIT.
- Anand, P. & Nevins, A. (2003). Shifty operators in changing contexts. *Proceedings of SALT XIV* 20-37.
- Baayen, R. H. (2008). Analyzing linguistic data: a practical introduction to statistics. Cambridge: Cambridge University Press.
- Banfield, A. (1982). Unspeakable sentences: narration and representation in the language of fiction. Boston, London, Melbourne, Henley: Routledge & Kegan Paul.
- Brener, R. (1983). Learning the deictic meaning of third person pronouns. Journal of Psycholinguistic Research 12(3), 235-62.
- Brocki, K. C. & Bohlin, G. (2004). Executive functions in children aged 6 to 13: a dimensional and developmental study. *Developmental Neuropsychology* **26**(2), 571–93.
- Charney, R. (1980). Speech roles and the development of personal pronouns. *Journal of Child Language* 7(3), 509-28.
- Clark, H. H. & Gerrig, R. J. (1990). Quotations as demonstrations. Language 66(4), 764-805.
- De Vries, M. (2008). The representation of language within language: a syntactico-pragmatic typology of direct speech. *Studia Linguistica* **62**(1), 39–77.
- Deutsch, W. & Pechmann, T. (1978). Ihr, dir, or mir? On the acquisition of pronouns in German children. *Cognition* **6**, 155–68.
- Diessel, H. (2012). Deixis and demonstratives. In C. Maienborn, K. von Heusinger & P. Portner (Eds.), *Semantics (HSK 33.3)*, vol. **III**, (pp. 1–25). Berlin, Boston: De Gruyter.

Ely, R. & McCabe, A. (1993). Remembered voices. Journal of Child Language 20(3), 671-96.

- Engberg-Pedersen, E. (1995). Point of view expressed through shifters. In K. Emmorey & J. S. Reilly (Eds.), *Language, gesture, and space* (pp. 133–54). Hillsdale, NJ: Lawrence Erlbaum Associates.
- 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**, 760–8.

- 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).
- Goodell, E. W. & Sachs, J. (1992). Direct and indirect speech in English-speaking children's retold narratives. *Discourse Processes* **15**(4), 395-422.
- 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. 203-28). Amsterdam: 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)* 1313-6.
- 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.
- Hollebrandse, B. (2007). A special case of wh-extraction in child language. *Lingua* **117**(11), 1897–906.
- Hothorn, T., Bretz, F. & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal* 50(3), 346–63.
- 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), 2017–36.
- Jansen, W., Gregory, M. L. & Brenier, J. M. (2001). Prosodic correlates of directly reported speech: evidence from conversational speech. In *Proceedings of the ISCA Workshop on Prosody in Speech Recognition and Understanding*, (pp. 77–80). Red Bank, NJ.
- 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.
- 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. (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, 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. *Journal of Language, Cognition and Neuroscience.* doi: 10.1080/23273798.2015.1047460.
- 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.
- Levinson, S. C. (2004). Deixis. In L. R. Horn & G. Ward (Eds.), *The handbook of pragmatics*, (pp. 97-121). Oxford: Blackwell.
- Lyons, J. (1977). Semantics, vol. 2. Cambridge: Cambridge University Press.
- 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. 133-45). Berlin, Heidelberg: Springer.
- Maier, E. (2015). Reported speech in the transition from orality to literacy. *Glotta*, **91**(1), 1–16.
- 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, DC: Georgetown University Press.
- Munro, R., Ludwig, R., Sauerland, U. & Fleck, D. W. (2012). Reported speech in Matses: perspective persistence and evidential narratives. *International Journal of American Linguistics* **78**(1), 41–75.

- Murphy, S. (1986). Children's comprehension of deictic categories in oral and written language. *Reading Research Quarterly* 21(2), 118-31.
- Nordqvist, Å. (2001). Speech about speech: a developmental study on form and function of direct and indirect speech. Göteborg: Kompendiet.
- Oliveira, M. & Cunha, D. A. C. (2004). Prosody as marker of direct reported speech boundary. In *Proceedings of Speech Prosody 2004* (pp. 263-6). Nara, Japan.
- Perner, J., Brandl, J. & Garnham, A. (2003). What is a perspective problem? Developmental issues in understanding belief and dual identity. *Facta Philosophica* 5, 355-78.
- 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) 1451-72.
- 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, 437–71.
- Quer, J. (2005). Context shift and indexical variables in sign languages. In E. Georgala & J. Howell (Eds.), *SALT XV* (pp. 152–68). Ithaca, NY: Cornell University.
- Recanati, F. (2000). Oratio Obliqua, Oratio Recta: an essay on metarepresentation. Cambridge, MA, London: MIT Press.
- Rice, K. D. (1986). Some remarks on direct and indirect discourse in Slave (Northern Athapaskan). In F. Coulmas (Ed.), *Direct and indirect speech* (pp. 47–76). Berlin, New York, Amsterdam: Mouton de Gruyter.
- Roberts, R. J. & Pennington, B. F. (1996). An interactive framework for examining prefrontal cognitive processes. *Developmental Neuropsychology* **12**(1), 105–26.
- Salazar Orvig, A., Marcos, H., Morgenstern, A., Hassan, R., Leber-Marin, J. & Parès, J. (2010). Dialogical beginnings of anaphora: the use of third person pronouns before the age of 3. *Journal of Pragmatics* 42(7), 1842–65.
- Schlenker, P. (2003). A plea for monsters. Linguistics and Philosophy 26(1), 29-120.
- 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. 1561–604). Berlin: Mouton de Gruyter.
- Shklovsky, K. & Sudo, Y. (2013). Indexical shifting in Uyghur and the syntax of monsters. In Seda Kan, Claire Moore-Cantwell, & Robert Staubs (Eds.), NELS 40: Proceedings of the 40th Annual Meeting of the North East Linguistic Society (pp. 151–166). Amherst, MA: GLSA, University of Massachusetts.
- Sullivan, K., Zaitchik, D. & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology* **30**(3), 395–402.
- 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.
- 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), 103–28.
- 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.
- Zwart, J.-W. (1997). Morphosyntax of verb movement: a Minimalist approach to the syntax of Dutch. Dordrecht: Kluwer.