

# Direct and Indirect Speech Interpretation by Older Adults

## (Bachelorproject)

Leonoor TB Ellen, s2409127, l.t.b.ellen@student.rug.nl,  
Jennifer Spenader\* and Emar Maier† and Franziska Köder‡

April 22, 2016

## Abstract

This study investigates the interpretation of direct speech (*Monkey said: "You get the ball"*) and indirect speech (*Monkey said that I get the ball*) in older adults (aged 65-95). Children interpret direct speech less accurately than indirect speech and recent work Köder and Maier (2016) has suggested that this may be because direct speech requires more working memory capacity and more inhibitory control. If this explanation is correct, we expect older adults, who are known to have fewer cognitive resources, to also show problems interpreting direct speech. 24 Dutch speaking older adults performed a referent-selection task, a Stroop task and a Digit Span task. We found that direct speech is interpreted less accurately than indirect speech by older adults. No relation with the Stroop task or the Digit span task was found. We discuss the results in comparison with earlier results of children (4;0-11;0) and young adults (18-61).

## 1 Introduction

When a speech act is reported, the whole sentence is referred to as *reported speech*. It constitutes a quote (*what* has been said), a person or character (*who* said it) and a speech act verb (Vincent and Perrin, 1999). The quote is the *reported* utterance and the whole sentence is the *reporting* utterance.

(1) John said to Mary: "I like your hair".

\*University of Groningen, Department of Artificial Intelligence

†University of Groningen, Department of Philosophy

‡University of Groningen, Department of Philosophy

(2) John said to Mary that he liked her hair.

In the first example above, the narrator reports literally what Sarah and John have said, which is called direct speech. Another option is a paraphrase, as is shown in the second example. Rather than literally reproducing the original words, as in direct speech, the quote is paraphrased and only the intention of the original speech act is reported. This is referred to as *indirect speech*.

### 1.1 Current study

In this study, a referent-selection paradigm designed by Köder and Maier (2016) is used to test the interpretation of direct and indirect speech in older adults (65+). Children interpret direct speech less accurately than indirect speech, while young adults interpret both equally well. The interpretation of direct and indirect speech might be influenced by inhibition and working memory. Since older adults show a decline in these functions, finding that older adults also show similar difficulties with direct speech would support the analysis that children's difficulties with direct speech are caused by cognitive limitations. Participants are presented with direct or indirect sentences in Dutch and are required to interpret the sentences by choosing the correct referent of deictic pronouns which are present in the sentence. The results indicate that older adults interpret direct speech less accurately than indirect speech, which is similar to children.

The difference between direct and indirect speech occurs both at a syntactic level and a semantic level.

## 1.2 Semantic difference

The semantic difference between direct speech and indirect speech concerns the perspective of the speaker. In a direct speech sentence, the original quote is spoken by the original speaker, and thus the perspective of the original speaker is reflected in the quote. In indirect speech, on the other hand, the sentence reflects the narrator's perspective, as the original quote is paraphrased (Coulmas, 2011). In the example given above, sentence (1) and (2) both describe the same situation. The reported speech in the first sentence, "I like your hair", reflects the point of view of John. He uses the pronoun "I" to refer to himself. In the second sentence, the original speech act is paraphrased and the word *that* is added. The second sentence reflects the point of view of narrator. In this case the pronoun *he* refers to John.

As the change in the example shows, the *pronoun* which is used in either sentence changes depending on the perspective of the speaker. These pronouns are referred to as *indexicals* (Roberts, 1993). In (1), the speaker is John. Thus, when he refers to himself, he says *I*. The person he is talking to is Mary. She is referred to as *you/your*. On the other hand, in (2), the narrator is 'talking'. John is now referred to as *he* and the possessive pronoun is changed to *her*. Thus, the interpretation of the pronouns signals the perspective of the sentence. Since sentences (1) and (2) have been written in direct speech and indirect speech respectively, direct speech reflects the perspective of the original speaker, while indirect speech reflects the perspective of the narrator.

Finally, the different perspective and their different pronouns also indicate different *contexts*. Direct speech refers to the context in which the reported speech was spoken originally, since it reflects the perspective of the original speaker. This context will be called the *original context*. Indirect speech, on the other hand, refers to the context in which the reporting speech was spoken. This is the *current context*, as it reflects the perspective of the (current) narrator.

In Dutch, direct and indirect speech have very different syntactic structures.

## 1.3 Indirect and direct speech in Dutch

In Dutch, direct and indirect speech are clearly distinguishable because of strong syntactical difference in word order and an obligatory subordinate clause marker, *dat* ('that'), as is shown in the following example:

- (3) Harry zei: "Ik heb de bal". (Harry said: "I have the ball".)
- (4) Harry zei dat hij de bal had (Harry said he had the ball.)

In sentence (4), the verb is moved to the end of the sentence and *dat* is added. The tense of the reported speech is the same as the tense of the original utterance in sentence (4), while it is not in sentence (3). This suggests that the distinction between direct and indirect speech are highly salient in Dutch and thus should not confuse the interpretation of the sentences.

The semantic and syntactic differences are reflected in the production and comprehension of indirect and direct speech, since direct and indirect speech are not produced in equal quantities.

## 1.4 Production and comprehension of reported speech

Direct speech is used more frequently than indirect speech by both children and adults. On the other hand, research on reported speech comprehension seems to indicate that indirect speech is interpreted more accurately than direct speech in certain situations.

Children produce more direct speech than indirect speech (Ely and McCabe, 1993), although the frequency of reported speech production in general increases when children grow older. On the other hand, Köder and Maier (2016) showed that children aged 4;0-11;0 perform worse at interpreting pronouns in direct speech than indirect speech in a referent selection task. In a different study, they argued that the perspective shift necessary for direct speech "increases the processing effort of deictic pronouns in speech reports" (Köder, Maier, and Hendriks, 2015), even in adults, which would explain the difficulty of direct speech for children.

Adults show the same patterns as children with

regard to production, yet not for comprehension. Vincent and Perrin (1999) found that direct speech is preferred over indirect speech by adults in oral discourse, 87% and 13% respectively. A referent selection task performed by Köder and Maier (2016) indicates that young adults do not show significant differences in comprehension between direct and indirect speech.

No research has been performed on older adults with regard to reported speech comprehension. However, in general, the decline of language comprehension and language production is disproportionate. While the accuracy of language production declines, resulting in more errors such as tip-of-the-tongue experiences, accuracy of language comprehension stays approximately the same with age (Burke and Shafto, 2004; Burke, Shafto, Craik, and Salthouse, 2008). In addition, Glosser and Deser (1992) show no difference between sentence comprehension of young and older adults at a syntactic level, which suggest that older adults should at least be able to distinguish direct and indirect speech based on its syntactic differences.

Köder and Maier (2016) suggest that one possible explanation for the difference in interpretation of the reported speech types in children is an insufficient working memory capacity and/or inhibitory control.

## 1.5 Working memory and inhibition

Working memory capacity, as introduced by Baddeley and Hitch (1974), is limited and short-term. It is crucial for both storing and manipulating information (Roberts and Pennington, 1996). It is well-known that children show increasing working memory capacity with age, while older adults show declines (Schneider, 2015; Waters and Caplan, 2001). If working memory capacity declines, less information can be stored and manipulated. Language, which is characterized by a temporal sequence of sounds or letters, depends on the working memory in the sense that the syntactical and semantic relations of words at the start of the sentence has to be remembered at the end of a sentence.

In order to decide the correct interpretation of pronouns in direct and indirect speech, the original context and the current context have to be taken into account. Because the original context occurs before the current context, its information has to be

stored in working memory. If working memory declines, however, the information of the original context will not be remembered as accurately, which could result in a less accurate interpretation of direct speech compared to indirect speech.

Hasher, Lustig, and Zacks (2007) argued that inhibition is important for cognitive functions. The main inhibitory function is to stop irrelevant information entering the consciousness. A Dutch study performed by Van der Elst, Van Boxtel, Van Breukelen, and Jolles (2006) suggest that the executive functions as measured by the Stroop task decline with age. For the interpretation of direct speech, the current context consists of irrelevant information, which has to be inhibited in favour of the relevant information. For indirect speech, the current information is correct and no distraction is present. Thus, indirect speech is expected to be more accurate than direct speech.

In sum, a decline in both functions would result in a less accurate interpretation of direct speech compared to indirect speech for older adults.

Our hypothesis is that indirect speech is easier than direct speech for older adults. Based on the decline in working memory and inhibitory functions, it is expected that older adults, similar to children, are less accurate in interpreting the correct reference of pronouns in direct speech than in indirect speech.

In addition to the main experiment, a Stroop task and a Working Memory span task will be performed to examine the relation between the inhibitory functions and working memory and the interpretation of direct and indirect speech.

## 2 Methods

### 2.1 Participants

The participants were 24 Dutch older adults between 65 and 95 ( $M = 75$ ,  $sd = 7.2$ ). Each participant volunteered to do the experiment. The data of one additional participant was removed due to medical reasons which impaired cognitive functions. In addition to the main task, all participants performed the Digit Span Task. 19 participants also performed the Stroop task.

## 2.2 Direct and indirect interpretation task

The main experiment is taken directly from Köder and Maier (2016). All sentences were in Dutch.

The participants were presented with three animals, of which all were voiced by a different male voice. All animals were male to prevent influence of the semantic difference between *she* and *he*. One of the animals was the leader. The leader determined to whom an object, also presented on the screen, was given. Only one animal could be the recipient of the object. The participant was required to select the animal to whom the object should be given, based on the uttered sentence. In addition to three practice items, the experiment consisted of two conditions. Both conditions are shown in Figure 1.

In the ‘no-report’ condition, the leading animal told the addressee, one of the other animals, to whom the object shall be given, audible to the participant:

- (5) Jij krijgt de bal ('You get the ball')

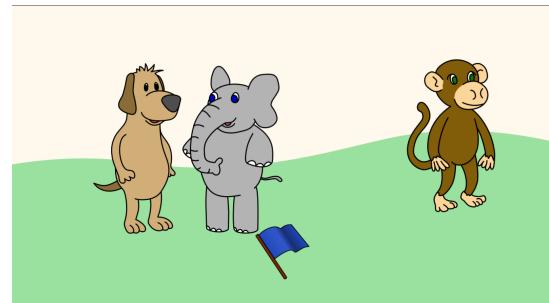
The participant then had to choose which animal would get the object.

The ‘report’ condition consisted of the reported speech sentences. The leading animal whispered them in the ear of a second animal, inaudible to the participant. This second animal reported what the leading animal has said to the third animal, either using direct speech or indirect speech:

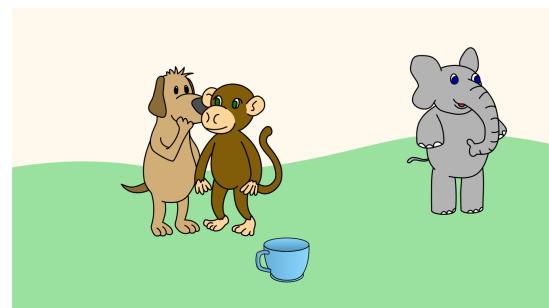
- (6) Hond zei: “Ik krijg de bal”. ('Dog said: "I get the ball")
- (7) Hond zei dat hij de bal krijgt ('Dog said that he gets the ball')

The speaker and the recipient of the object alternated between the trials. Consequently, the pronoun used, either *ik* ('I'), *jij* ('you') and *hij* ('he'), in the reporting utterance varied for each sentence. The direct and indirect speech items, the three pronouns and the speaker of the sentences were presented randomly. The ‘no-report’ condition consisted of 15 trials with 5 trials for each pronoun, and the ‘report’ condition consisted of 15 trials for each sentence type, with 5 trials for each pronoun. Thus, the participant was presented with 45 items in total.

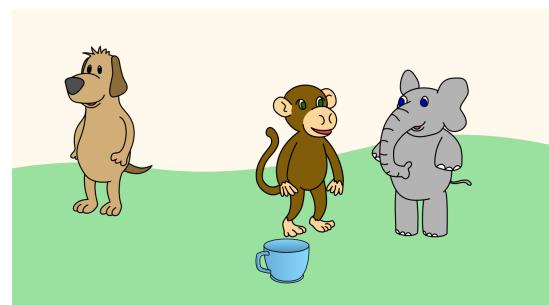
Before the ‘no-report’ and ‘report’ conditions,



(a) No report condition. Dog tells Elephant who gets the flag, for example: “Jij krijgt de vlag” (You get the flag).



(b) Report condition part 1. Dog whispers the sentence in the ear of Monkey, inaudible to the participant. Monkey subsequently reports what Dog has said to Elephant.



(c) Report condition part 2. Monkey reports what Dog has said to Elephant, audible to the participant. The reported sentence is either in direct or indirect speech.

Figure 1: Conditions. (a) depicts the ‘no-report’ condition. (b) and (c) depict the ‘report’ condition.

the participant was presented with two kinds of practice items. When the animal was introduced, they were required to select the correct animal when asked the question ‘Who is animal X?’. This

tested if the participant understood which animal was which. Secondly, one of the animals told the participant who gets the object with the use of the name of the animal:

- (8) Olifant krijgt het boek ('Elephant gets the book')

This ensured that participant understood the instructions. If a participant chose the wrong animal in either case, they were given feedback and the item was repeated until the answer was correct.

The experiment was conducted on a computer either with a mouse or touch screen. If the participants were capable of handling a mouse, they could select the preferred option by clicking on the animal. If the participants were not able to use a mouse, the researcher clicked on the right answer after it was given by the participant. In two cases, a touch screen was used, so the participants could press the screen with their finger. Finally, the main experiment was performed together with two other language-related experiments, in addition to the Stroop and Digit Span task. The order in which the experiments were conducted was randomized.

The accuracy and choice of the participants was recorded for each condition. The results of the older adults will be compared to the results obtained by Köder and Maier (2016).

### 2.3 Stroop task

ROOD  ROOD

Figure 2: Items of the Stroop task for each card, A (left), B (middle) and C (right).

The Stroop task which was used is a Dutch version (Hammes, 1973) of the original task created by Stroop (1935) and is intended to measure inhibition. Three cards were presented to the participants. Each contained a ten by ten grid of items. The first card (A) showed the words *blauw*, *rood*, *groen* and *geel* (blue, red, green and yellow) in black ink. The second card (B) showed those same colours in small rectangles in ink of that colour. The third card (C) showed the same words, but printed in incongruent combinations of word name and colour. For example, the word *rood* ('red') is printed with

blue ink. See Figure 2 for an example. The participants were required to name the ink colour, and not the word. Thus, the correct responses to each item from left to right in Figure 2 are 'rood' (red), 'rood' and 'blauw' (blue) respectively. No congruent items (matching word meaning and word ink colour) were included. The time needed to say all the items aloud and the number of errors the participants made were recorded. The interference time was computed by subtracting card B, the colour naming time, from card C, the incongruent reading time). This calculation is suggested to relate most closely to the inhibitory effect (Jensen, 1965). This interference time was used in the analysis.

### 2.4 Forward-backward digit span task

The working memory test consisted of a forward-backward digit span task. In both cases, multiple lists of digits were recited to the participant with a one-second interval and no list pronunciation (raise in voice signalling the end of a list). The number of digits in the lists differed. Two items were presented for each number. If the participants failed to repeat both of the two items correctly, the test was stopped and the score was calculated by counting the correct items. If the participants repeated both items correctly, the number was increased by one, starting with 2 and going up to 9 for the forward digit span and up to 8 for the backward digit span.

In the forward digit span test the participants were asked to repeat the list of digits as is. In the backward digit span test the participants were asked to repeat the digits backwards. In addition, the participants were given two practice items of two digits each. The final score was calculated by adding the score for the forward and the backward test and comparing this score to a standard score table.

The added total score is used in the analysis.

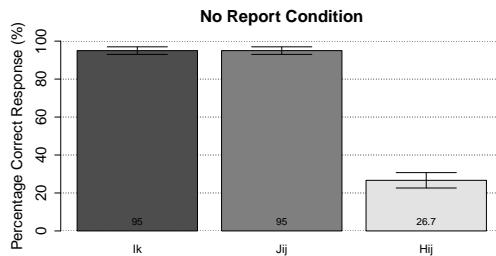
## 3 Results

All results of this experiment were compared with the results obtained by Köder and Maier (2016).

We first present the results of the interpretation of the 'no-report' condition and the results of the interpretation of direct and indirect speech items

for the older adults alone. Then, we compare the older adults with the children and young adults.

### 3.1 ‘No-report’ condition



**Figure 3: Mean and standard error of the ‘no report’ condition**

Figure 3 shows the mean and standard error for the three different pronoun types. Pronoun *hij* had a low accuracy of 27.6 % ( $se = 4.4\%$ ). The other two pronouns, *ik* and *jij* had a high accuracy of 95% ( $se = 1.8\%$ ) and 96% ( $se = 2.1\%$ ) respectively.

The logistic mixed effect model, created with a stepwise variable addition procedure, showed a fixed effect for Pronoun type ( $p < 0.001$ ) and random intercept for each subject ( $p < 0.001$ ).

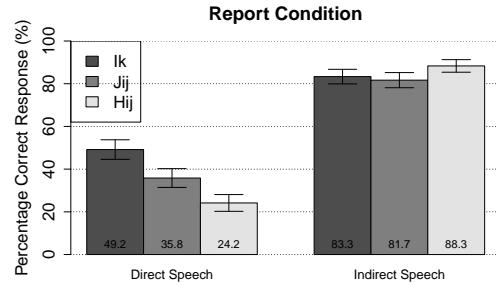
We compared the means of different combinations of Pronoun type with multiple comparisons (Tukeys contrast) from the ‘multcomp’ package, version 1.4-1. *Hij* differed from both *ik* ( $p < 0.001$ ) and *jij* ( $p < 0.001$ ). It was interpreted less accurately than both. *Jij* and *ik* did not differ significantly ( $p = 0.49$ ).

Neither the Digit Span score nor the Stroop Interference time contributed to the model.

### 3.2 Report condition

Figure 4 shows the mean and standard error for each pronoun and report type.

A logistic linear mixed-effect regression model was created with a stepwise variable addition procedure without Stroop Interference time. This model consisted of a significant *fixed*-effect for Pronoun ( $p = 0.023$ ), Speech Type ( $p < 0.001$ ) and an interaction effect between Pronoun and Speech Type ( $p = 0.002$ ). In addition, it contained a random effect of Participant ( $p = 0.003$ ) with a random slope per



**Figure 4: Mean percentage of correct pronoun interpretation and standard error of the ‘report’ condition.**

Speech Type for each participant ( $p < 0.001$ ). The procedure was repeated, this time with all participants who did not complete the Stroop task omitted. The second model consisted of a *fixed*-effect for Speech Type ( $p < 0.001$ ) and Pronoun ( $p = 0.051$ ). In addition, the model contained a random slope for Speech Type ( $p < 0.001$ ) per participant. Pronoun was included in the model because of an significant interaction effect between Speech Type and Pronoun ( $p = 0.003$ ). Thus, both models contained the same variables. Neither the Stroop Interference time nor Digit Span score added significantly to the model.

Subsequent results are based on the model with all participants included. The indirect speech sentences in the model created differed significantly from direct speech sentences ( $p < 0.001$ ). The interpretation of direct speech is less accurate than the interpretation of indirect speech.

Multiple comparisons (Tukeys contrast) were done for the interaction between Pronoun and Speech Type. The results are shown in Table 1. There are no significant differences among the indirect speech items, thus neither one is interpreted more accurately than the other. Pronoun *hij* in the direct speech items differs significantly from all other Pronouns in both sentences except for *jij* in the direct speech sentences ( $p < 0.001$  and  $p = 0.127$ , respectively). It has the lowest accuracy of all items. Pronoun *ik* in the direct speech sentence differs from *hij* and *ik* in the indirect speech sentence ( $p = 0.006$  and  $p = 0.03$ , respectively). *Jij* in the direct speech condition differs from all three pronouns in the indirect speech condition (*hij*:  $p$

**Table 1: Comparison pronouns in direct and indirect speech with mutliplecomparison of mean (Tukey Contrast) for accuracy of Pronoun interpretation**

Linear Hypotheses:		Estimate	Std. Error	z-value	Pr(> z )	
Direct <i>ik</i>	<i>vs</i> Direct <i>hij</i> = 0	1.7184	0.3658	4.697	< 0.001	***
Direct <i>jij</i>	<i>vs</i> Direct <i>hij</i> = 0	0.8442	0.3543	2.383	0.12661	
Direct <i>jij</i>	<i>vs</i> Direct <i>ik</i> = 0	-0.8742	0.3382	-2.585	0.07637 .	
Indirect <i>ik</i>	<i>vs</i> Indirect <i>hij</i> = 0	-0.6763	0.4829	-1.400	0.67552	
Indirect <i>jij</i>	<i>vs</i> Indirect <i>hij</i> = 0	-0.8657	0.4779	-1.812	0.39621	
Indirect <i>jij</i>	<i>vs</i> Indirect <i>ik</i> = 0	-0.1894	0.4357	-0.435	0.99736	
Indirect <i>ik</i>	<i>vs</i> Direct <i>ik</i> = 0	3.0919	1.0757	2.874	0.03391	*
Indirect <i>jij</i>	<i>vs</i> Direct <i>jij</i> = 0	3.7766	1.0734	3.518	0.00402	**
Indirect <i>hij</i>	<i>vs</i> Direct <i>hij</i> = 0	5.4866	1.1157	4.918	< 0.001	***
Indirect <i>ik</i>	<i>vs</i> Direct <i>hij</i> = 0	4.8103	1.0873	4.424	< 0.001	***
Indirect <i>jij</i>	<i>vs</i> Direct <i>hij</i> = 0	4.6208	1.0808	4.275	< 0.001	***
Indirect <i>hij</i>	<i>vs</i> Direct <i>ik</i> = 0	3.7682	1.1044	3.412	0.00593	**
Indirect <i>jij</i>	<i>vs</i> Direct <i>ik</i> = 0	2.9025	1.0691	2.715	0.05372 .	
Indirect <i>hij</i>	<i>vs</i> Direct <i>jij</i> = 0	4.6424	1.1085	4.188	< 0.001	***
Indirect <i>ik</i>	<i>vs</i> Direct <i>jij</i> = 0	3.9661	1.0800	3.672	0.00241	**

$< 0.001$ , *ik*:  $p = 0.002$ , *jij*:  $p = 0.004$ ). Both are interpreted less accurately than the others.

In summary, direct speech was interpreted less accurately than indirect speech. Pronoun *hij* in indirect speech was interpreted the least accurate of all pronouns.

Finally, the errors were analysed by interpreting the choices of the participants. 97% of the incorrectly interpreted direct speech sentences were interpreted as if the sentence was indirect.

### 3.3 Comparison children and adults

The results of he older adults were compared with the results of the children and young adults gained by Köder and Maier (2016). First, the ‘no-report’ conditions will be analysed. Secondly, the ‘report’ conditions will be analysed.

#### 3.3.1 Comparison ‘no-report’ condition

Figure 5 shows the ‘no-report’ condition for each of the age groups (Child, Adult and Senior).

The logistic linear mixed-effect regression model consisted of a *fixed*-effect for Pronoun ( $p < 0.001$ ) and Age Group ( $p < 0.001$ ) and an interaction effect between Pronoun and Age Group ( $p < 0.001$ ). The interaction between Pronoun and Age Group was analysed using the `testInteraction` function

from the ‘phia’-package (version 0.2-0) in R. Pronoun *hij* differed significantly between age groups Adult and Child ( $p = <0.001$ ). Adults interpreted the pronoun more accurately than children. The difference of *hij* between Child and Senior had a low significance level of  $p=0.08$ . Children are less accurate than Seniors. Pronoun *jij* differed between age groups Child and Senior ( $p = 0.045$ ), with Seniors being less accurate. Finally, pronoun *ik* did not differ between age groups.

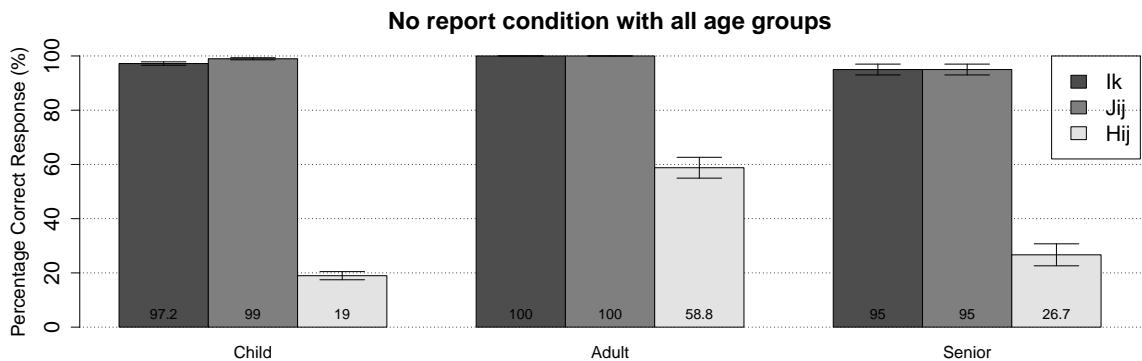
Thus, the only difference between the ‘no-report’ conditions of each age group is the interpretation of *hij*, which is interpreted less accurately by children than by adults. Older adults neither show a difference between children nor adults. However, the mean suggests that they are less accurate than young adults, but more accurate than children.

#### 3.3.2 Comparison ‘report’ condition

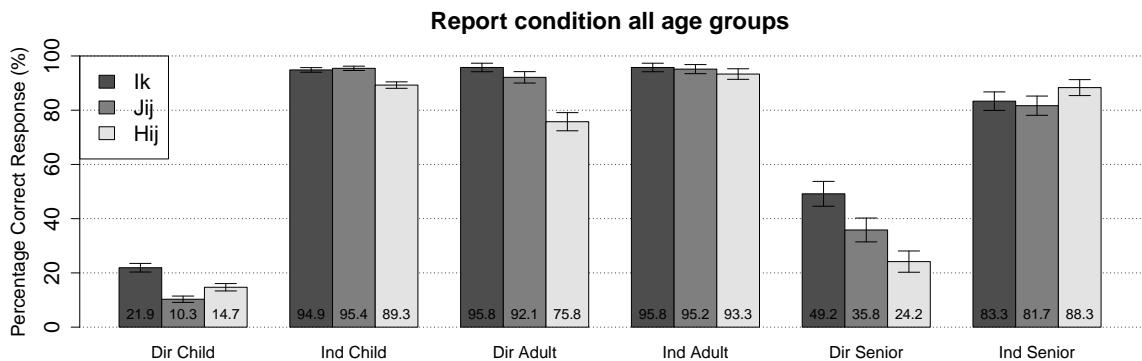
Figure 6 shows the results for the ‘report’ condition.

The model consisted of *fixed*-effects for Pronoun, Speech Type and Age Group, as well as a random slope for Speech Type per participant. In addition, an interaction effect between Pronoun, Speech Type and Age Group was included.

Table 2 shows the comparison between Speech Type for each age group. For indirect speech, all



**Figure 5:** Mean percentage of pronoun interpretation per age group of the ‘no-report’ condition. Error bars indicate error.



**Figure 6:** Mean percentage of pronoun interpretation of the ‘report’ condition for each age group. Error bars indicate standard error.

age groups did not show a significant difference. For direct speech, all age groups showed a significant difference. Children are the least accurate in their interpretation of direct speech. Older adults are more accurate than children, but less accurate than young adults. The young adults are the most accurate.

The full interaction between Speech Type, Pronoun and Age Group was analysed with the `testInteraction`-function in the ‘phia’-package (version 0.2-0) with ‘Holm’ p-value adjustment. The results are shown in Table 3. In direct speech, the accuracy for pronoun *hij* differed between Adults and Child and between Adult and Senior, but not between Child and Senior. Children and

older adults interpreted *hij* less accurately than young adults. The accuracy of pronoun *ik* in the direct speech sentences was significantly different between each age group. Children were the least accurate, young adults the most accurate and older adults in between. Similarly, the accuracy of *jij* was significantly different between each age group, in the same order. In indirect speech, the accuracy of the pronoun *ik* differed for Child and Senior, but not between Child and Adults or between Adults and Seniors. Seniors are the least accurate for *ik* in the indirect speech sentences. The accuracy of pronoun *jij* in the indirect speech items differs between Children and Senior, but again not between Children and Adults or between Adults and Seniors.

**Table 2: Comparison Direct and Indirect speech per Age Group**

Sentence Type	Age Group	Value	Df	Chisq	Pr(>Chisq)	
Dir	<b>Adult vs Child</b>	0.99852	1	132.0755	< 0.001	***
Dir	<b>Adult vs Senior</b>	0.99114	1	46.7002	< 0.001	***
Dir	<b>Child vs Senior</b>	0.14193	1	11.7337	0.002	**
Ind	Adult vs Child	0.61634	1	0.5357	0.464	
Ind	Adult vs Senior	0.88056	1	5.6999	0.051	.
Ind	Child vs Senior	0.82108	1	5.4796	0.051	.

Older adults are the least accurate. All other pairs do not differ significantly.

### 3.3.3 Children aged 11 and older adults

In addition, in order to analyse the relation between the age groups, a second model was created. The age group Child was divided into children and 11-year-olds. The model included the same variables as the previous model. Figure 7 shows the direct speech interpretation of each age group. The interaction is shown in Table 4. 11-year-olds differ significantly from both Child and Adults in the direct speech sentences, but don't differ significantly from Seniors. Younger children are the least accurate, 11-year-olds and older adults are equally accurate and young adults are the most accurate in interpreting direct speech.

## 3.4 Additional analysis

The incorrect choices of the *hij* pronoun in the ‘no-report’ condition was analysed. Out of all incorrect choices, 94 % was due to the interpretation of *hij* as referring to the addressee rather than to the animal outside of the discourse.

The interpretation of the incorrect pronouns in the ‘report’ condition. 97% of all incorrectly interpreted pronouns in the direct speech sentences were interpreted as if they were situated in an indirect speech sentence. A similar pattern was found by Köder and Maier (2016) for children: 99% of the incorrect interpreted direct speech pronouns were interpreted as if they were indirect speech.

## 4 Discussion

Older adults clearly interpret direct speech less accurately than indirect speech (36.4% and 86.1% re-

spectively). Thus, they differ from young adults and show a pattern similar to the pattern shown in children. This suggest that older adults have a similar difficulty as do children with interpreting direct speech in this task. Older adults do not differ from 11-year-olds.

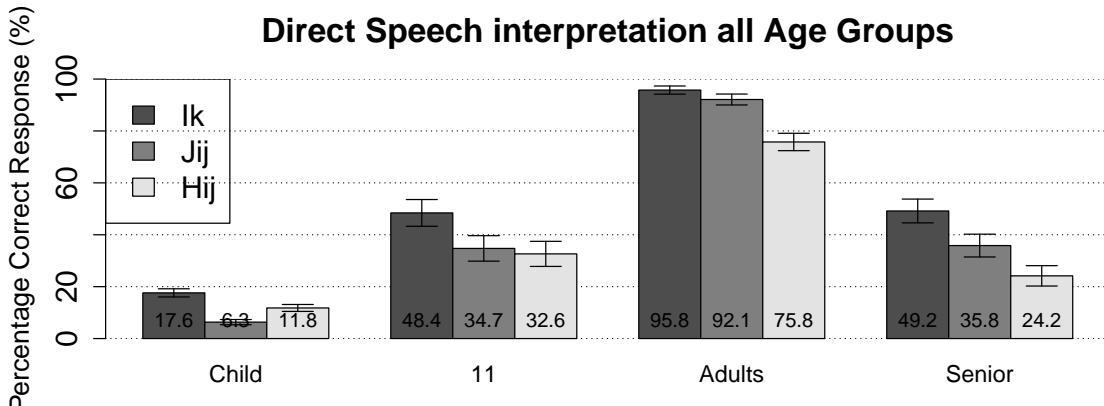
Köder and Maier (2016) propose that the difference between direct and indirect speech is caused by the difficulty of the perspective shift between the two report types. The perspective of direct speech is assumed to be more demanding, since the listener is required to shift to the original event. Köder et al. (2015) found evidence supporting this view in young adults. The decision time for pronouns *you* and *he* and the number of mistakes was significantly higher for direct speech than indirect speech. Children, then, would lack the resources needed to perform the perspective shift. In a similar fashion, as cognitive functions of older adults decline, the increase in resources for direct speech would affect their accuracy. Their inability to perform the perspective shift might be supported by the fact that 97% of the incorrectly interpreted direct speech sentences were interpreted as if they were indirect speech.

In addition, the difference might be caused by the different functions of direct and indirect speech. Köder et al. (2015) argue that the paradigm of this experiment favours indirect speech because of the nature of the task. The main purpose of the reported speech is the transfer of information, while direct speech is used predominantly in a narrative context (Vincent and Perrin, 1999). In this task, information is being transferred, namely the animal who will receive the object. The context is not a narrative. Thus, indirect speech is favoured.

Neither the Stroop task nor the Digit span task were significant, thus no explicit link between inhibition, Working Memory Capacity and the dif-

**Table 3: Results comparison Age Group, Pronoun type and Speech Type.**

Pronoun	Age Group	Value	Df	Chisq	Pr(>Chisq)	
DIR	ik Adult vs Child	1.00	1	92.93	<0.001	***
	ik Adult vs Senior	0.99	1	36.68	<0.001	***
	ik Child vs Senior	0.12	1	12.43	<0.001	**
	jij Adult vs Child	1.00	1	128.06	<0.001	***
	jij Adult vs Senior	0.99	1	42.31	<0.001	***
	jij Child vs Senior	0.08	1	18.14	<0.001	***
	hij Adult vs Child	0.99	1	74.31	<0.001	***
	hij Adult vs Senior	0.98	1	28.79	<0.001	***
	hij Child vs Senior	0.28	1	2.55	0.66	
IND	ik Adult vs Child	0.61	1	0.30	1.00	
	ik Adult vs Senior	0.92	1	6.55	0.075	
	ik Child vs Senior	0.88	1	7.96	0.043	*
	jij Adult vs Child	0.51	1	0.00	1.00	
	jij Adult vs Senior	0.92	1	6.76	0.075	
	jij Child vs Senior	0.92	1	11.32	0.008	**
	hij Adult vs Child	0.72	1	1.87	0.856	
	hij Adult vs Senior	0.74	1	1.30	1.00	
	hij Child vs Senior	0.52	1	0.02	1.00	



**Figure 7: Mean percentage of pronoun interpretation in direct speech for age groups Children, 11-year-old, Adults and Seniors. Error bars indicate standard error.**

ference between direct and indirect speech comprehension has been found. However, the Stroop task was only performed by 19 participants, all aged 65-95. A larger number of participants, including young adults, might be a better reflection of the population and might enlarge any possible results. Furthermore, the Stroop task as performed for this study might not be suitable for more refined con-

clusions. It did not take the number of errors into account. Using another inhibitory task might shine more light on the influence on inhibition. According to Burke, MacKay, and James (2000), the Stroop task in general might not be a sufficient measurement for inhibition, since other non-inhibitory explanations have been proposed.

The Digit Span task also did not influence ac-

**Table 4: Interaction including 11-year-olds.**

Age groups	Value	Df	Chisq	Pr(>Chisq)	
DIR	11 vs Adults	0.01	1	39.65	<0.001 ***
	11 vs Child	0.91	1	16.80	<0.001 ***
	11 vs Senior	0.55	1	0.07	1.000
	Adult vs Child	1.00	1	149.39	<0.001 ***
	Adult vs Senior	0.99	1	49.37	<0.001 ***
	Child vs Senior	0.11	1	17.95	<0.001 ***
IND	11 vs Adult	0.39	1	0.20	1.000
	11 vs Child	0.51	1	0.00	1.000
	11 vs Senior	0.83	1	2.49	0.575
	Adult vs Child	0.61	1	0.49	1.000
	Adult vs Senior	0.88	1	5.53	0.131
	Child vs Senior	0.82	1	5.23	0.133

curacy. Similarly to the Stroop task, a different, more extensive method could be used to test working memory. Increasing the number of participants would also create a more accurate representation. However, based on these results, no conclusion can be drawn about the involvement of working memory in the task.

The result of the ‘no-report’ condition show a clear difference between pronoun types. *Hij* is interpreted less accurate than *jij* and *ik*. As expected based on the results by Köder and Maier (2016), children and young adults differ: children perform less accurate than young adults. For older adults, the accuracy of *hij* does not differ significantly from either children or adults, thus hard conclusions can not be drawn.

The difference in accuracy of the *hij* pronoun in the ‘no-report’ condition might be explained by the relation between the position of the animals and the perspective of the participant. The experiment assumes that the participant is not part of the discourse, but rather an observer. However, the participant might interpret the experiment differently. If the participant is part of the spoken discourse, the interpretation of the second and third pronoun changes. The animal can be interpreted as talking to the participant, instead of the (supposed) addressee. *Jij* would then refer to the participant. However, the participants themselves are not available options and this interpretation would probably correctly discarded. The interpretation of the third pronoun is influenced by the position of the speaking animal. The speaking animal is turned towards

one of the other animals, which can be interpreted as a deictic marker. *Hij*, then, would refer to the addressee (for reference see Figure 1a), as the position ‘points’ towards this animal. This is supported by the fact that 94% of the incorrect interpretation of *hij*, the addressee is chosen. In indirect speech, such a problem can not occur. The referent of *hij* in the sentence ‘Aap zei dat hij de bal krijgt’ (Monkey said that he gets the ball) is specified by the whispering sequence, in which it is clear who is speaking to whom. Thus, the participant can not be part of the discourse in indirect speech. A variation of the task, in which the participants themselves is a valid option, could indicate if the participants indeed interpret the discourse as involving themselves.

In conclusion, the interpretation of direct speech and indirect speech of older adults is similar to children aged 11. They interpret direct speech less accurately than indirect speech, which is in contrast to the production of direct and indirect speech, which favours direct speech. Whether this result is caused by the nature of the task, the difficulty in the perspective shift or something else has yet to be determined. However, it is clear that reported speech interpretation changes with age.

## References

- Alan D Baddeley and Graham J Hitch. Working memory. *The psychology of learning and motivation*, 8:47–89, 1974.
- Deborah M Burke and Meredith A Shafto. Aging

- and language production. *Current directions in psychological science*, 13(1):21–24, 2004.
- Deborah M. Burke, Donald G. MacKay, and Lori E. James. *Theoretical approaches to language and aging*, pages 204–237. Oxford University Press, New York, NY, US, 2000.
- Deborah M Burke, Meredith A Shafto, FIM Craik, and TA Salthouse. Language and aging. *The handbook of aging and cognition*, 3:373–443, 2008.
- Florian Coulmas. *Reported Speech: Some general issues*, pages 1–29. Walter de Gruyter, Meuchen, DEU, 2011.
- Richard Ely and Allyssa McCabe. Remembered voices. *Journal of Child Language*, 20:671–696, 10 1993. ISSN 1469-7602.
- Guila Glosser and Toni Deser. A comparison of changes in macrolinguistic and microlinguistic aspects of discourse production in normal aging. *Journal of Gerontology: PSYCHOLOGICAL SCIENCES*, 47:266–272, 1992.
- J. Hammes. *De Stroop Kleur-woord Test: Handleiding [The Stroop Color-Word Test: Manual]*. Swets & Zeitlinger, Amsterdam, 1973.
- Lynn Hasher, Cindy Lustig, and RT Zacks. Inhibitory mechanisms and the control of attention. *Variation in working memory*, pages 227–249, 2007.
- Arthur R Jensen. Scoring the stroop test. *Acta Psychologica*, 24:398–408, 1965.
- Franziska Köder and Emar Maier. Children mix direct and indirect speech: evidence from pronoun comprehension. *Journal of Child Language*, FirstView:1–24, 2 2016. ISSN 1469-7602.
- Franziska Köder, Emar Maier, and Petra Hendriks. Perspective shift increases processing effort of pronouns: a comparison between direct and indirect speech. *Language, Cognition and Neuroscience*, 30(8):940–946, 2015.
- Lawrence D. Roberts. *How Reference Works : Explanatory Models for Indexicals, Descriptions, and Opacity*. SUNY Series, Scientific Studies in Natural and Artificial Intelligence. State University of New York Press, 1993. ISBN 9780791415764.
- RJ Roberts and BF Pennington. An interactive framework for examining prefrontal cognitive processes. *Developmental neuropsychology*, 12(1):105–126, 1996. PT: J; UT: WOS:A1996UM97000007.
- Wolfgang Schneider. *Memory Development from Early Childhood Through Emerging Adulthood*, chapter The Role of Basic Memory Capacities and Working Memory, pages 131–181. Springer International Publishing, Cham, 2015.
- J Ridley Stroop. Studies of interference in serial verbal reactions. *Journal of experimental psychology*, 18(6):643, 1935.
- Wim Van der Elst, Martin P. J. Van Boxtel, Gerard J. P. Van Breukelen, and Jelle Jolles. The stroop color-word test: Influence of age, sex, and education; and normative data for a large sample across the adult age range. *Assessment*, 13 (1):62–79, 2006.
- Diane Vincent and Laurent Perrin. On the narrative vs non-narrative functions of reported speech: A socio-pragmatic study. *Journal of Sociolinguistics*, 3(3):291–313, 1999.
- Gloria S Waters and David Caplan. Age, working memory, and on-line syntactic processing in sentence comprehension. *Psychology and aging*, 16 (1):128, 2001.

## A Appendix

**Table A1: Older Adults ‘No-report’ condition: Model = Accuracy ~ Pronoun + (1 | Subject)**

Predictor	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	4.0289	0.6582	6.121	< 0.001 ***
$J_{ij}$	-0.0000	0.6198	-0.000	1.000
$H_{ij}$	-5.5781	0.6961	-8.013	< 0.001 ***

**Table A2: Older Adults ‘Report’ condition: Model = Accuracy ~ Pronoun \* SpeechType + (SpeechType | Subject)**

Predictor	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-0.0737	0.4481	-0.164	0.869
$J_{ij}$	-0.8742	0.3382	-2.585	< 0.05 **
$H_{ij}$	-1.7184	0.3658	-4.697	< 0.001 ***
IndSpeech	3.0919	1.0756	2.874	< 0.05 **
$J_{ij}:\text{IndSpeech}$	0.6848	0.5516	1.241	0.214
$H_{ij}:\text{IndSpeech}$	2.3947	0.6058	3.953	< 0.001 ***

**Table A3: Comparison Age Groups ‘No Report’ condition: Model = Accuracy ~ Pronoun \* Age-Group + (1 | Subject)**

Predictor	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	1.0809	1.0186	1.061	0.289
$I_k$	22.5378	23.2313	0.970	0.332
$J_{ij}$	31.2065	19.9976	1.561	0.119
Child	-6.0304	1.2341	-4.886	< 0.001 ***
Senior	-2.9657	1.4600	-2.031	< 0.05 **
$I_k:\text{Child}$	-10.5731	23.2304	-0.455	0.649
$J_{ij}:\text{Child}$	-17.7552	19.9986	-0.888	0.375
$I_k:\text{Senior}$	-15.7100	23.2442	-0.676	0.499
$J_{ij}:\text{Senior}$	-24.3787	20.0120	-1.218	0.223

**Table A4: Comparison Age Groups ‘Report’ condition: Model = Accuracy ~ Pronoun \* SpeechType \* AgeGroup + (SpeechType | Subject)**

Predictor	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	5.0858	0.6846	7.429	< 0.001	***
$J_{ij}$	-1.0068	0.6034	-1.669	0.095	
$H_{ij}$	-3.1721	0.5855	-5.418	< 0.001	***
IndSpeech	0.3102	1.1019	0.282	0.778	
Child	-7.1787	0.7433	-9.658	< 0.001	***
Senior	-5.1938	0.8566	-6.063	< 0.001	***
$J_{ij}:\text{IndSpeech}$	0.8014	0.8793	0.911	0.362	
$H_{ij}:\text{IndSpeech}$	2.4542	0.8435	2.910	< 0.05	**
$J_{ij}:\text{Child}$	-0.4133	0.6360	-0.650	0.516	
$H_{ij}:\text{Child}$	2.4008	0.6113	3.927	< 0.001	***
$J_{ij}:\text{Senior}$	0.0796	0.6958	0.114	0.909	
$H_{ij}:\text{Senior}$	1.3513	0.6918	1.953	0.051	
IndSpeech:Child	6.7485	1.1735	5.751	< 0.001	***
IndSpeech:Senior	2.7236	1.4397	1.892	0.059	
$J_{ij}:\text{IndSpeech}:\text{Child}$	0.8200	0.9561	0.858	0.391	
$H_{ij}:\text{IndSpeech}:\text{Child}$	-2.9389	0.9033	-3.254	< 0.05	**
$J_{ij}:\text{IndSpeech}:\text{Senior}$	-0.0649	1.0412	-0.062	0.950	
$H_{ij}:\text{IndSpeech}:\text{Senior}$	0.0516	1.0408	0.050	0.960	

**Table A5: Age Groups inclusive 11-year-olds ‘Report’ condition: Model = Accuracy ~ Pronoun \* SpeechType \* AgeGroup + (SpeechType | Subject)**

Predictor	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	0.0046	0.5591	0.008	0.993	
$J_{ij}$	-1.1305	0.4289	-2.636	< 0.05	**
$H_{ij}$	-1.3192	0.4355	-3.029	< 0.05	**
IndSpeech	5.3011	1.2951	4.093	< 0.001	***
Adults	4.9442	0.8667	5.705	< 0.001	***
Child	-2.3982	0.6149	-3.900	< 0.001	***
Senior	-0.0972	0.7384	-0.132	0.895	
$J_{ij}:\text{IndSpeech}$	0.7828	0.9360	0.836	0.403	
$H_{ij}:\text{IndSpeech}$	-0.1808	0.9273	-0.195	0.845	
$J_{ij}:\text{Adults}$	0.1400	0.7355	0.190	0.849	
$H_{ij}:\text{Adults}$	-1.7942	0.7208	-2.489	< 0.05	**
$J_{ij}:\text{Child}$	-0.4417	0.4903	-0.901	0.368	
$H_{ij}:\text{Child}$	0.6716	0.4777	1.406	0.160	
$J_{ij}:\text{Senior}$	0.2156	0.5492	0.393	0.695	
$H_{ij}:\text{Senior}$	-0.4795	0.5698	-0.841	0.400	
IndSpeech:Adults	-4.8186	1.6306	-2.955	< 0.05	**
IndSpeech:Child	2.0559	1.3488	1.524	0.127	
IndSpeech:Senior	-2.2489	1.5812	-1.422	0.155	
$J_{ij}:\text{IndSpeech}:\text{Adults}$	0.0022	1.2819	0.002	0.999	
$H_{ij}:\text{IndSpeech}:\text{Adults}$	2.5762	1.2476	2.065	< 0.05	**
$J_{ij}:\text{IndSpeech}:\text{Child}$	1.0833	1.0250	1.057	0.291	
$H_{ij}:\text{IndSpeech}:\text{Child}$	-0.3901	0.9884	-0.395	0.693	
$J_{ij}:\text{IndSpeech}:\text{Senior}$	-0.0594	1.0883	-0.055	0.956	
$H_{ij}:\text{IndSpeech}:\text{Senior}$	2.6669	1.1106	2.401	< 0.05	**

**Table A6: Items.** A fixed set of pronouns and sentence types are created by combining the objects and animals randomly.

Objects (O)	Animals (A)	Example sentences
de auto (the car)	Olifant (Elephant)	“Ik krijg de auto” (I get the car)
het boek (the book)	Aap (Monkey)	
de gitaar (the guitar)	Hond (Dog)	
de hoed (the hat)		“Aap zei dat hij de gitaar krijgt” (Monkey said that he gets the guitar)
de kop (the cup)	Pronouns (P)	
de lepel (the spoon)	Ik (I)	
de paraplu (the umbrella)	Jij (you)	
de pen (the pen)	Hij (He)	
de roos (the rose)		“Hond zei: ‘Jij krijgt de lepel’ ” (Dog said: “You get the spoon”)
de schaar (the scissors)	Sentence Types	
de sjaal (the scarf)	P krijg(t) O	
de tandenborstel (the toothbrush)	“A zei dat P O krijg(t)”	
de vlag (the flag)	“A zei: ‘P krijg(t) O’ ”	
het vliegtuig (the plane)		“Olifant zei: ‘Ik krijgt de paraplu’ ” (Elephant said: “I get the umbrella”)
de voetbal (the ball)		
de zonnebril (the sunglasses)		