

The repeated name penalty effect in children's natural reading: Evidence from eye tracking

Sarah Eilers, Simon P Tiffin-Richards
and Sascha Schroeder

Quarterly Journal of Experimental
Psychology
1–10
© Experimental Psychology Society 2018
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1747021818757712
qjep.sagepub.com



Abstract

We report data from an eye tracking experiment on the repeated name penalty effect in 9-year-old children and young adults. The repeated name penalty effect is informative for the study of children's reading because it allows conclusions about children's ability to direct attention to discourse-level processing cues during reading. We presented children and adults simple three-sentence stories with a single referent, which was referred to by an anaphor—either a pronoun or a repeated name—downstream in the text. The anaphor was either near or far from the antecedent. We found a repeated name penalty effect in early processing for children as well as adults, suggesting that beginning readers are already susceptible to discourse-level expectations of anaphora during reading. Furthermore, children's reading was more influenced by the distance of anaphor and antecedent than adults', which we attribute to differences in reading fluency and the resulting cognitive load during reading.

Keywords

Eye tracking; children's reading; repeated name penalty effect; anaphora processing

Received: 10 January 2017; revised: 12 November 2017; accepted: 5 December 2017

Introduction

Reading comprehension processes have been characterised as an online effort to build a mental model of what a text is about. Abstracted from the text surface, readers integrate the content of each new proposition into the existing set of propositions in an incremental way (van Dijk & Kintsch, 1983; Zwaan, Langston, & Graesser, 1995; Zwaan & Radvansky, 1998). Readers therefore need to evaluate how every new proposition fits into the mental model that they have constructed thus far. Protagonists, in particular, serve as anchors for the global coherence of a text. The repeated reference to the same set of discourse entities has been termed “referential continuity” and constitutes a powerful coherence marker (Garnham, Oakhill, & Johnson-Laird, 1982; Givón, 1983). This is important because referential continuity is signalled explicitly on the text level by the use of anaphora; wherever there is a continuity of referents in the text, this is explicitly marked by anaphoric expressions. The fact that readers are sensitive to the fit of referring expressions in the discourse is demonstrated by the repeated name penalty (RNP) effect. In

this study, we investigated children's sensitivity to the form of referential expressions in a discourse. The RNP is informative for the study of children's reading because it allows conclusion about children's ability to direct attention to discourse-level processing cues during reading. More precisely, we studied similarities and differences in children's and adults' processing of repeated names and pronouns when they are either near to their antecedent or far from their antecedent in a three-sentence story. Using an eye tracker, we recorded readers' eye movements to obtain a detailed picture of developmental differences in the time course of reading processes when resolving pronouns and repeated names.

Max Planck Institute for Human Development, Berlin, Germany

Corresponding author:

Sarah Eilers, MPRG Reading Education and Development (REaD), Max Planck Institute for Human Development, Lentzeallee 94, D-14195 Berlin, Germany.
Email: eilers@posteo.de

The RNP effect

The RNP effect is a well-established finding in the literature on anaphoric processing. The term “repeated name penalty” was coined in the seminal paper by Gordon, Grosz, and Gilliom (1993). In their Experiment 1, they measured young adults’ reading times for passages containing a proper name in the subject position of the first sentence, and either pronouns or repeated names in the subject position of three subsequent sentences. They observed decreased reading times for sentences containing pronouns compared with repeated names, while comprehension accuracy was comparable across conditions. The authors interpret their findings within the Centering theory framework (Grosz, Joshi, & Weinstein, 1983; Grosz, Weinstein, & Joshi, 1995). Centering theory establishes a set of formal rules about the appropriate anaphor for an entity based on its relative prominence in the discourse. Following Centering theory, the most prominent entity in the discourse should be pronominalised. The experimental literature has shown repeatedly that adult readers prefer pronouns as a referent for prominent discourse entities (e.g., Fukumura & van Gompel, 2015).

An alternative explanation for the RNP effect is offered by the Informational Load Hypothesis (ILH; Almor, 1999). The ILH explains the effect in terms of memory interference between representation of the referent in the current situation model and the representation of the referential expression (Almor & Nair, 2007; Peters, Boiteau, & Almor, 2016). Repeated names are semantically rich, as they carry substantive information and associations. Pronouns, in contrast, are semantically uninformative and code only number and gender. Importantly, such an uninformative anaphor is expected when a referent is highly accessible in the discourse, for example, because the referent was mentioned recently or is the only available discourse entity at present (Ariel, 2001; Kehler, 2002). The general idea behind the ILH was already formulated in Grice’s cooperative principle of quantity, that is, to make a contribution as informative as required, but not more informative than required (Grice, 1975). Names are typically used to introduce new referents, which is inconsistent with their use as referring expressions for prominent discourse entities. In other words, repeated names in a phrase where a pronoun could be used clash with readers’ discourse model and impede sentence processing because readers assume some added value of the repeated name and spend time trying to integrate superfluous information (Almor & Nair, 2007). In summary, the ILH stresses cognitive access to the referent, whereas Centering theory concentrates on the linguistic features of the discourse to determine the type of anaphor used. However, both make identical predictions regarding the RNP in adults; a repeated name slows discourse processing compared with a pronoun when it refers to a prominent, or accessible,

discourse entity. Our aim in this study was to elicit the RNP effect in children using eye tracking measures and compare it with the effect in adults.

Self-paced reading time studies have repeatedly shown that for salient discourse entities, reading times increase when a repeated name is used instead of a pronoun (see Lezama, 2015, for a review). Using eye tracking, it has further been shown that in contexts where a pronoun is expected, repeated names increase the likelihood of regressive eye movements (Kennison & Gordon, 1997). The RNP may even have effects on aspects of reading comprehension but only in highly skilled readers who make use of the pronoun as a local coherence marker (Shapiro & Milkes, 2004).

More generally, the implication of the RNP is that proficient readers are sensitive to the type of referring expression during reading processing because the type of referring expression chosen in the text is directly linked to discourse coherence. Proficient readers evaluate incoming information not only on the text surface level but also from a discourse representational point of view. The main aim of the current experiment was to investigate the RNP effect in beginning readers because it is currently unclear whether beginning readers use discourse context in a similar way to skilled adult readers.

Development of the RNP effect

For children, the RNP has been studied online in listening comprehension (Megherbi & Ehrlich, 2009). Engelen, Bouwmeester, de Bruin, and Zwaan (2014) studied 6- to 11-year-old children’s eye movements in a visual world paradigm while they were listening to a complex story involving multiple characters. The probability that children fixated the target increased after the mention of a proper name, but not a pronoun. They also assessed comprehension of these texts and found that good comprehenders were more likely to make anticipatory eye movements to the referent of a pronoun than poor comprehenders. The authors discuss the possibility that poor comprehenders lack the ability to make inferences during listening of complex stories. This has been suggested before in studies which assessed comprehension. When complex inferences are required to identify the referent of a pronoun, poor comprehenders among the children failed to name the correct referent in up to one-third of the presented items (Oakhill & Yuill, 1986). It can be concluded that referential processing could be facilitated for children when inferences are not required because of a text-level identity between antecedent and referential expression. This is particularly relevant for children’s reading. Children read more slowly and spend more time on single words than adults (Blythe & Joseph, 2011). Their reading can be characterised as more effortful, associated with an increase in cognitive load. The

establishing prominence in the discourse may be difficult for children, and they may not use discourse-level cues in the building of situation models as efficiently as adults. As the RNP effect has been explained in terms of the accessibility of referents in working memory, we predicted that children should not show the same RNP as adults. On the contrary, beginning readers' processing downstream from the repeated name might be facilitated because when a repeated name is used, antecedent and anaphor can be mapped directly at the word level (Gernsbacher, 1989). Children may profit from the repeated name because the referential expression and referent are identical on the text surface level, unlike the pronoun, which requires a local inference. Pronouns require inferences that span several words in the text, which arguably poses a challenge for beginning readers. We know of only one eye tracking study that focused specifically on children's comprehension of anaphora during reading. Joseph, Bremner, Liversedge, and Nation (2015) had children read short paragraphs in which (a) the distance between antecedent and anaphor and (b) the semantic typicality of the antecedent (typical: a truck—the vehicle, atypical: a crane—the vehicle) were manipulated. Although the authors did not find effects of distance in early online measures, children did make more regressions out of far than near anaphors in their study. We are further aware of two studies which directly tested the RNP effect in children's reading. Ehrlich, Remond, and Tardieu (1999) conducted a self-paced reading study with 10-year-old children. The children read expository texts with a repeated noun phrase or a pronoun in subject position. The RNP effect was not a main focus in their study, but the authors report elevated reading times for sentences with a pronoun rather than a repeated name. Interestingly, this was particularly true for skilled readers. However, the texts in their study were rather complex and therefore, it may have been particularly difficult for the children to resolve the pronoun in these texts. In a study with 10-year-old German children, Schimke (2015) compared reading times of the verb downstream from the anaphor in sentences such as "John is sitting in the ground and John/he/Ø draws a picture," where the second noun phrase was a noun, a pronoun, or an ellipsis (i.e., it was omitted entirely, which is possible in German). She found a penalty for the repeated name compared with the elliptical subject, but not the pronoun. These results do not directly compare with the existing findings of the RNP in English reviewed above. It is not clear how the discourse integration of a repeated name or a pronoun relates to the discourse integration of an ellipsis. Taken together, previous studies into children's reading of pronouns and repeated names suggest that children do show some sensitivity to Anaphor Type. However, these studies did not directly compare adults' and children's reading, so developmental differences have not yet been addressed.

The current study

In this article, we report a natural reading experiment with children using a repeated name manipulation of short, three-sentence stories. Children and adults read these paragraphs while their eye movements were recorded. We used a two-factorial design, contrasting pronouns and repeated names in three-sentence discourse contexts where the anaphor was either near or far from the antecedent. We included a distance manipulation based on findings of a prior eye tracking study on children's processing of anaphora during natural reading (Joseph et al., 2015). If it is the case that cognitive load is a relevant factor for children's processing of pronouns and repeated names, we argue that distance of anaphor and antecedent should play a role for the RNP in children. This is because a direct mapping of anaphor and antecedent may be particularly helpful for beginning readers when they are further apart in the story.

The RNP effect typically spills over to the region following the anaphor, which was identical in all conditions in our reading materials. We were therefore particularly interested in two regions: The anaphor itself (anaphor region) and the region directly following the anaphor (post-anaphor region). We analysed first fixation times and gaze durations to tap into early processing effects in the post-anaphor region. Furthermore, we analysed total reading times to pick up later effects of Anaphor Type in the post-anaphor region.

We expected to find differences in the way a repeated name affects adults' and children's reading processing which lead to two different sets of hypotheses for children and adults. For the adults, we expected to replicate the RNP effect with our materials using eye tracking. In line with the well-established RNP literature, we hypothesised a processing advantage for pronouns over repeated names. We expected longer first fixation times and gaze durations in the post-anaphor region following a repeated name than following a pronoun. Second, building on previous findings (Kennison & Gordon, 1997), we expected adults to make more regressions from repeated names than pronouns in the anaphor region. Because the items were short and written for a primary school reading-level, we did not expect that distance to the antecedent would induce any difficulty for the adults.

For the children, in contrast to the adults, we hypothesised a processing advantage for repeated names over pronouns. We hypothesised shorter first fixation times and gaze durations after a repeated name than a pronoun in the post-anaphor region. This would suggest that children rely more on surface-level text information during reading (mapping of information) and do not use discourse-level cues for online situation model building as efficiently as adults do. Second, in contrast to the adults, we did not expect children to make more regressions out of repeated names than pronouns. If the

pronoun is more difficult for the children to integrate than the repeated name, there should however be more regressions from the post-anaphor region following a pronoun than a repeated name. This hypothesis is the opposite of our expectations for the adults. In addition, for the children, we predicted longer gaze durations in the post-anaphor region after distant pronouns than near pronouns because of the added difficulty of connecting lexical information that spans longer distances of text (Joseph et al., 2015). Finally, if we found an interaction of Anaphor Type and Distance to the antecedent for reading time measures in children, we would expect the repeated name to ease processing of distant anaphors, indicated by shorter first fixation times and gaze durations. Such a finding would imply that distance to the antecedent affects the processing of pronouns and repeated names differentially in children, such that the pronoun is even more difficult to resolve when the antecedent is further away in the text. We did not expect such an interaction between Anaphor Type and Distance for the adults.

Method

Participants

We recruited 29 fourth graders from three Berlin schools who took part in two sessions. From these, 23 full datasets were obtained. Five children were excluded because of missing data due to technical issues, and the data from one child were excluded because he had learned German after the age of 6 years. Of the remaining 23 children, 9 were girls. In addition, 25 native German-speaking adults were recruited via university mailing lists.

The children were 9 to 10 years old ($M=9$ years, standard deviation (SD)=15 months). The adults were $M=25.2$ years old, $SD=38.5$ months, and 17 were women. All participants reported normal or corrected-to-normal vision. The participants completed a standardised reading fluency test (SLRT-II; Moll & Landerl, 2010). Children did not differ from the population mean in either word reading fluency, $M=53.0$, $SD=24.4$, $t(22)<1$, $p=.56$, or non-word reading fluency, $M=55.5$, $SD=22.6$, $t(22)=1.2$, $p=.25$. Our adult sample did not differ from the population mean in word reading fluency, $M=51.1$, $SD=29.9$, $t(24)<1$, $p=.86$, but was slightly above average in non-word reading fluency, $M=63.0$, $SD=29.4$, $t(24)=2.13$, $p=.04$. Children additionally completed a standardised reading comprehension test (ELFE; Lenhard & Schneider, 2006). Importantly, our sample did not differ significantly from the population mean on either the word comprehension subscale, $M=-0.07$, $SD=0.8$, $t(22)<1$, $p=.68$, or the text comprehension subscale, $M=-0.15$, $SD=1.0$, $t(22)<1$, $p=.46$.

Materials

Items. Materials consisted of 52 three-sentence stories. For each of the 52 items, four different stimulus versions were created in which the factors Anaphor Type (repeated name vs pronoun) and Distance (near vs far) were manipulated in a within-item design. The stories comprised 16 to 17 words (89-108 characters) and were structurally similar. The introductory sentence of each story contained a referent and an activity of the referent. The structure of the stories is demonstrated in Table 1.

The target sentence with our main regions of interest contained either a personal pronoun (pro) or a repeated name (rpn) to refer to the referent. All target sentences were of the form adverb—verb—subject (anaphor region)—direct object (post-anaphor region). The anaphor was in the middle of each target sentence, which corresponds to standard word order in German. The sentences were kept simple and contained age-appropriate topics for fourth graders. Word frequencies of the direct object in the post-anaphor region were derived from the German children's book corpus childLex (Schroeder, Würzner, Heister, Geyken, & Kliegl, 2015). The mean normalised lemma frequency of the direct objects was high, $M=89.4$, $SD=104.2$.

One extra sentence, designed to lengthen the distance between antecedent and anaphor, appeared either in the middle of the story or in final position. It never introduced a referent that could be confused with the target referent and was plausible within the story in both positions. Prior to the eye tracking experiment, we had 40 children of the same age group as our child sample (mean age $M=9.7$, $SD=.54$, 20 of them girls) rate the items for plausibility and difficulty. The children took part in a 45-min paper-pencil group session, including breaks, at their school. Children were asked to read the stories silently and afterwards rate them on a 4-point scale, where 1=*very implausible/very difficult to read* and 4=*very plausible/very easy to read*. Our manipulations did not affect plausibility or difficulty: An analysis of variance (ANOVA) with the factors Anaphor and Distance yielded no significant effects for either comprehension or plausibility, all $F(1, 204)<1$.

Apparatus. We used an EyeLink 1000 eye tracker (SR Research, Ontario, Canada) to record eye movements during reading at a rate of 1,000Hz. The stories were presented on an ASUS LCD monitor (21") with a refresh rate of 120Hz. The stories appeared in the middle of the screen in a 4:3 frame. The stories were presented using SR Research Experiment Builder. All stories appeared continuously in two to three lines, in Courier New, font size 16, using black letters on a white background. Although line breaks occasionally occurred before or after a region of interest, this was then the case for all

Table 1. Structure of stimulus materials.

Condition	Story	Anaphor	Distance
Introduction (invariable)	Peter steigt aus dem Bett. Peter gets up from his bed.		
1	Sofort macht [Peter] [das Frühstück.] Right away, Peter prepares breakfast. Der Rest der Familie schläft noch friedlich. The rest of the family is still fast asleep.	rpn	near
2	Der Rest der Familie schläft noch friedlich. The rest of the family is still fast asleep. Sofort macht [Peter] [das Frühstück.] Right away, Peter prepares breakfast.	rpn	far
3	Sofort macht [er] [das Frühstück.] Right away, he prepares breakfast. Der Rest der Familie schläft noch friedlich. The rest of the family is still fast asleep.	pro	near
4	Der Rest der Familie schläft noch friedlich. The rest of the family is still fast asleep. Sofort macht [er] [das Frühstück.] Right away, he prepares breakfast.	pro	far

The anaphor is written in bold face, and square brackets indicate the regions of interest used in analyses. English translations (non-literal) are printed in grey.

conditions of that item. Participants were seated at a monitor distance of 62 cm in a head-and-chin rest. Recording of the eyes was monocular and only the left eye was tracked.

Procedure. Written informed consent was obtained from the children's parents ahead of the study, and oral consent was obtained from each child prior to testing. Adult participants signed an informed consent form. For each item, one of the story versions was assigned to one of four item lists according to a Latin square design. Participants were assigned to one of the lists based on their order of appearance. Children took part in two sessions at their school. The paper-pencil part of the test was administered in one group session. The individual sessions were conducted in a quiet room that was suitable for eye tracking provided by the school. Adults were tested in the facilities of the Max Planck Institute for Human Development in Berlin.

A 5-point calibration was conducted for each participant until calibration error reached a maximum of 0.5° of visual angle. After the first calibration, participants read three practice stories, each with a following comprehension question. They were instructed to read the stories silently, press a button on a gamepad after having finished reading and answer the comprehension question via button-press. Comprehension questions appeared randomly after 25% of trials. The questions never tapped comprehension of the pronoun but were designed to ensure attentive reading, for example, "Was the family wide awake?" (see Table 1).

Analysis. Data were inspected and y-axis drift corrections were applied as necessary using the DataViewer software (SR Research, version 1.11.9). Fixations were cleaned automatically using the DataViewer four-stage fixation cleaning: At Stage 1, fixations shorter than 80 ms and within 0.5° from the neighbouring fixation were merged with each other. At Stage 2, fixations shorter than 40 ms and within 1.25° distance were merged with a neighbouring fixation. At Stage 3, all interest areas were checked for at least three neighbouring fixations of less than 140 ms and if found, these were merged. At Stage 4, only fixations between 120 and 1,200 ms (for children data) and between 80 and 1,000 ms (for adult data) were kept. The cleaning removed about 13% of fixations of the children, and about 16% of fixations of the adults. Finally, before models for the dependent measures were calculated for each eye movement measure, all observations above 2.5 standard deviations from the person or item mean of each dependent measure were deleted from the fixation record (roughly 2% of observations).

We calculated four eye tracking measures for the anaphor and the post-anaphor region: *first fixation time* (duration of the first fixation that falls into the area of interest), *gaze duration* (summed duration of first-pass fixations), *total reading time* (summed fixations in a region), and *regression probability* (the likelihood of a leftward saccade out of a region).

Reading time data were analysed with linear mixed-effects models and regression probability was analysed with generalised linear mixed-effects models, using the lme4 package version 1.7 (Bates, Maechler, & Bolker,

Table 2. Results of mixed-effects models.

	First fixation time		Gaze duration		Total reading time		Regression probability	
	Anaphor	Anaphor+I	Anaphor	Anaphor+I	Anaphor	Anaphor+I	Anaphor	Anaphor+I
Distance	0.10	12.45***	9.80**	15.00***	0.05	17.31***	9.64**	325.75***
Anaphor	1.62	0.42	121.02***	13.87***	165.10***	1.16	12.14***	13.24***
Age	38.66***	12.80***	71.32***	73.57***	48.75***	71.13***	3.49	3.88*
Distance × Anaphor	3.83	0.07	2.21	1.12	0.02	0.07	0.51	1.88
Distance × Age	3.38	3.94*	9.46**	0.01	3.13	0.53	6.66**	89.49***
Anaphor × Age	0.01	0.53	74.08***	0.86	59.19***	0.29	1.16	0.62
Distance × Anaphor × Age	0.51	1.89	1.35	2.57	0.20	1.33	2.31	1.90

ANOVA: Analysis of variance. F-values for first fixation time, gaze duration and total reading time. χ^2 values for regression probability. * $p < .05$; ** $p < .01$; and *** $p < .001$.

2012) in *R* (R Development Core Team, 2016). We calculated individual models for each region of interest and each dependent variable with Anaphor (repeated name vs pronoun), Distance (near vs far), and Age (child vs adult) as fixed effects, and participants and items as crossed random intercepts.

All reading time measures were log-transformed to achieve a more normal distribution. To ease interpretation, the back-transformed results are reported in milliseconds. The significance of the fixed effects was determined using effects coding and type-II model comparisons in the ANOVA function in the car package (Fox, Friendly, & Weisberg, 2013). Post hoc comparisons were estimated using cell-means coding and single-degree-of-freedom contrasts as implemented in the glht function in the multcomp package (Hothorn et al., 2015).

Results

Global measures

Mean comprehension accuracy for the adults was high, $M=97%$, $SD=18%$, and slightly lower for children, $M=92%$, $SD=27%$, but consistently above chance level. Adults and children differed in mean text reading time, which amounted to an averaged $M=8.7$ s, $SD=4.2$ s, for children, whereas adults took $M=3.9$ s, $SD=1.6$ s, to read the stories. Consequently, we found a large effect of Age group for all our dependent reading time measures (see Table 2). As children's reading is characterised by more and longer fixations compared with adults' (Blythe & Joseph, 2011), this was to be expected and we will concentrate on interactions of Age with Anaphor and Distance in the remainder of this article.

Regions of interest

We will report our results by region, starting with the post-anaphor region. Note that we will not report effects of Anaphor for reading time measures in the anaphor region

itself because these cannot be separated from word length and frequency of pronouns (short, frequent) and names (longer, less frequent). The model means for all dependent measures can be found in Table 3.

Post-anaphor region. There was no effect of Anaphor for first fixation time. For gaze duration, however, there was a main effect of Anaphor such that regions following repeated names took longer to read, $M=527$ ms, standard error (SE)=25 ms, than regions following pronouns, $M=484$ ms, $SE=23$ ms. We found no interaction of Anaphor and Age in gaze duration, indicating an RNP effect for both adults and children. For regression probability, we found a main effect of Anaphor such that both groups were more likely to make regressions out of the post-anaphor region following pronouns, $M=0.37$, $SE=0.03$, than repeated names, $M=0.29$, $SE=0.03$.

There was further a main effect of Distance in the post-anaphor region, such that gaze durations were longer after anaphora that were far from the antecedent, $M=528$ ms, $SE=25$ ms, than those that were near their antecedent, $M=484$ ms, $SE=23$ ms. There was no interaction of Distance and Anaphor in gaze duration. Furthermore, there were early main effects of Distance and an interaction of Distance and Age in first fixation times. Planned contrasts revealed that children showed longer first fixation times in the post-anaphor region after far anaphors, $M=216$ ms, $SE=5$ ms, than after near anaphors, $M=201$ ms, $SE=5$ ms, $t=-3.89$, $p<.001$. In contrast, the effect of Distance for first fixation time in adults was not significant, $t=-1.14$, $p=.25$. Furthermore, there was no effect of Anaphor but a significant main effect of Distance for total reading time: The post-anaphor region following far anaphors was read for longer, $M=682$ ms, $SE=35$ ms, than following near anaphors, $M=635$ ms, $SE=32$ ms. We further found a large main effect of Distance and an interaction of Distance and Age for regression probability. Although the Distance effect was significant for both adults, $t=-19.5$, $p<.001$, and children, $t=-8.9$, $p<.001$, post hoc contrasts showed that it was significantly larger for the adults compared with the children, $t=9.5$, $p<.001$.

Table 3. Model means for dependent measures.

	First fixation time		Gaze duration		Total reading time		Regression probability	
	Anaphor	Post-anaphor	Anaphor	Post-anaphor	Anaphor	Post-anaphor	Anaphor	Post-anaphor
Adults								
pro								
far	185 (5)	195 (5)	194 (9)	363 (23)	213 (12)	447 (31)	.09 (.02)	.77 (.04)
near	183 (5)	187 (5)	193 (9)	336 (21)	204 (12)	427 (29)	.04 (.01)	.12 (.02)
rpn								
far	192 (5)	188 (5)	198 (9)	390 (25)	229 (13)	466 (32)	.12 (.02)	.71 (.04)
near	183 (5)	188 (5)	198 (9)	356 (23)	225 (12)	433 (30)	.06 (.01)	.09 (.02)
Children								
pro								
far	217 (5)	214 (6)	253 (11)	725 (48)	290 (17)	1,028 (73)	.07 (.01)	.47 (.05)
near	229 (6)	203 (5)	267 (12)	623 (41)	302 (17)	922 (66)	.09 (.02)	.26 (.04)
rpn								
far	228 (6)	218 (6)	336 (15)	757 (50)	424 (24)	1,011 (72)	.16 (.02)	.42 (.05)
near	225 (6)	200 (5)	389 (17)	736 (49)	437 (25)	953 (68)	.12 (.02)	.14 (.03)

Back-transformed, rounded model means for the dependent measures in the anaphor and post-anaphor regions of interest. Standard errors are given in parentheses.

Two additional analyses were conducted to investigate the unexpected effect of Distance. First, we analysed regression probability for the region in paragraph-final position, which was the post-anaphor region in the far condition and the last region of the extra sentence in the near condition, respectively. We included Distance and Age as fixed effects, and subjects and items as random effects in a generalised linear mixed-effects model with regression probability as the dependent variable. There was a main effect of Age, $t=22.8$, $p<.001$, but no effect of Distance, $t<1$, and importantly no interaction of Distance and Age, $t=1.8$, $p=.175$. In the final region of the paragraph, regression probability was higher for adults, $M=80\%$, $SE=4\%$, than children, $M=47\%$, $SE=6\%$.

Second, we conducted a similar analysis including first-pass skipping probability on all words preceding the post-anaphor region as dependent variable. The main effects of Distance and Age were not significant, both $t<1$, but there was an interaction of Distance and Age, $t=15.4$, $p<.001$, such that adults' first-pass skipping probability was higher in the far condition when there was more text between anaphor and referent, $M=0.08$, $SE=0.01$, than in the near condition, $M=0.07$, $SE=0.01$, $t=-2.7$, $p<.01$, whereas for children, it was higher in the near condition, $M=0.08$, $SE=0.01$, than in the far condition, $M=0.06$, $SE=0.01$, $t=2.9$, $p<.01$.

These two complementary analyses suggest two differences in the overall patterns of eye movements between children and adults which were unrelated to the anaphor manipulation. In contrast to children, adults showed a greater increase in regression probability in the final region of the paragraph regardless of where the anaphor was positioned, as well as slightly increased

skipping probability when more text was between referent and anaphor.

Anaphor region. There was a main effect of Anaphor for regression probability in the anaphor region, such that regressions were more likely from repeated names, $M=11\%$, $SE=1\%$, than pronouns, $M=7\%$, $SE=0.1\%$. There was no interaction of Anaphor and Age for regression probability, which suggests that the effect of Anaphor was the same for adults and children.

We further found effects of Distance in the anaphor region. For gaze durations, planned contrasts revealed that Distance had an effect on children's gaze durations, $t=4.21$, $p<.0001$, but not adults', $t<1$, $p=.90$. Children read near anaphors, $M=322$ ms, $SE=13$ ms, slower than far anaphors, $M=291$ ms, $SE=12$ ms, independent from Anaphor Type. There were no effects of Distance, nor interactions of Distance and Age, in first fixation time or total reading time.

For regression probability, we found a main effect of Distance and an interaction of Distance with Age. Post hoc contrasts revealed that the Distance effect in regression probability was driven solely by the adults. There was a simple main effect of Distance in adults, $t=-4.0$, $p<.001$, but not in children, $t<1$, $p=.77$, such that adults made more regressions when the anaphor was far from its antecedent than when it was close to the antecedent.

Finally, higher skipping rates for pronouns than repeated names were expected in the anaphor region, as pronouns are often skipped during reading (e.g., Drieghe, Desmet, & Brysbaert, 2007). We found that pronouns were skipped more often than repeated names by adults and children.

Adults had a skipping rate of 40% ($SD=49\%$) for pronouns and 21% ($SD=41\%$) for repeated names. Children skipped considerably less, with a skipping rate of 16% ($SD=37\%$) for pronouns and 4% ($SD=19\%$) for repeated names. A generalised linear mixed-effects model over skipping rate in the anaphor region showed significant main effects of Anaphor, $t=51.0$, $p<.001$, and Age, $t=18.0$, $p<.001$.

Discussion

We conducted an eye tracking experiment to compare the RNP effect in children's and adults' natural reading. Both groups read short, three-sentence stories with a single, salient discourse entity introduced in the first sentence. We used a two-factorial design varying Anaphor Type (repeated name vs pronoun) and Distance to the antecedent (near vs far). The aim was to replicate the RNP for adults using eye tracking and contrast adults' and children's sentence processing. Our hypotheses for the adults based on the existing literature were that they show longer reading times after a repeated name than a pronoun, and more regressions from repeated names than pronouns. For the children, in contrast, we expected to see longer reading times for regions following pronouns than repeated names because pronouns add the necessity to connect information across several words in the paragraph for local inference. We expected children to benefit from a repetition of text surface information during reading processing. Surprisingly, we saw more similarities than differences in children's and adults' processing of the paragraphs, and conclude that 9-year-old children already show sensitivity to discourse-level information during text reading.

The RNP effect manifested in children and adults as longer gaze durations in the post-anaphor region after the repeated name compared with after the pronoun, indicating increased integration difficulty when the name is repeated (Lezama, 2015). We observed more regressions directly out of repeated names than pronouns for both age groups. In line with effects reported by Kennison and Gordon (1997), the repeated name induced more regressions as soon as it was encountered. This suggests that adults and children initiate immediate repair strategies when faced with unexpected information at the discourse-level. Importantly, our results suggest that children, despite the fact that their reading is much slower and more effortful than adults', already anticipate the appropriate form of discourse referent during reading. Both age groups seem to expect a pronoun when the referent is salient in the text, and their processing is disrupted when a repeated name is used, that is, when textual information clashes with discourse expectation. Note that since we are comparing a pronoun with a repeated name, the regression behaviour in the anaphor region may to some extent be driven by lexical characteristics of the anaphor. In research designs comparing pronouns and names, it is not

possible to control for lexical features. Length and frequency are not the only lexical differences between pronouns and names. As we have discussed in the introduction, the ILH essentially assumes that the RNP is a result of the semantic richness of content words (repeated name) compared with function words (pronoun). Put differently, the ILH predicts that the processing difference for pronouns and repeated names follows from lexical characteristics, which are a defining feature of the two types of anaphor.

Although distance to the antecedent clearly had a detrimental effect on anaphor processing, distance did not modulate the RNP. With only one referent in the discourse, the extra sentence between antecedent and anaphor may not have introduced sufficient intervening linguistic material to licence the use of a repeated name. Presumably, readers still anticipated a recurrent entity and the appropriate pronominal reference. Further studies may want to investigate whether children and adults process a repeated name similarly in contexts where there is more than one referent present.

We expected an interaction of Anaphor Type and Distance to the antecedent for the children and no effect of distance for the adults. Instead, we found that Distance to the antecedent influenced processing independently from Anaphor Type for adults and children. First, we found an early effect of distance to the anaphor for children's first fixation time in the post-anaphor region. Given that the lexical content of the post-anaphor region does not differ between the conditions, and given that the effect is too early to reflect integration effort in the post-anaphor region, we interpret it as a spillover effect from the anaphor region. Longer first fixation durations in children may reflect an increased processing load for referring expressions when their antecedent is further away. Converging effects emerged for both age groups in gaze duration and total reading time, which may indicate that adults and children need more effort to integrate anaphors that are far from their antecedent. The fact that we did not find an interaction of Distance to the antecedent and Anaphor Type in the two early measures suggests that it is not pronoun resolution or the repeated name in particular, but more generally the integration of a distant referring expression which leads to delayed processing in readers of both age groups. In line with this interpretation, total reading times were longer in the post-anaphor region when these were far from their antecedents. This suggests that children and adults take more time to integrate anaphors when these are further away from their antecedents. This finding is generally consistent with prior work on the processing of near and distant typical and atypical anaphors in children (Joseph et al., 2015).

We found that both age groups made more regressions from the post-anaphor region when it is further away from its antecedent. This distance effect was stronger for the adults than the children, which is unexpected and contradicts our initial hypothesis. Because the Distance effect did not

interact with Anaphor Type, we assume that it is independent from Anaphor Type and therefore not attributable to the RNP. The fact that the effect was stronger for adults than children and surfaced in late processing measures puts its connection to processing difficulty into question. Two options will be considered here which may explain the unexpected effect of the distance between referent and anaphor in regression probability. It has been shown that in skilled readers, regressions are more likely from sentence-final regions than mid-sentence regions (Rayner et al., 2000). Note that in the far condition, the whole paragraph has been presented when readers reach the post-anaphor region and therefore the likelihood to regress to earlier sections may be greater. The results from the exploratory analysis of regression probability from final regions suggest that the position of the region in the paragraph is a critical determining factor for regression probability in adults. Moreover, in the far condition, adults were more likely to skip a word on first-pass and revisit it in a second pass. The high skipping rate of the anaphor region for adults in particular may be explained by the high predictability of the referring expression in the given discourse (e.g., Drieghe, Brysbaert, Desmet, & De Baecke, 2004). Finally, there was a main effect of Anaphor Type on regression probability in the post-anaphor region which went in the opposite direction from the effect in the anaphor region. This finding directly contradicts our hypotheses, as we have predicted more regressions after the repeated name than the pronoun. Note, however, that we found more regressions immediately in the anaphor region, which we have interpreted in terms of an RNP. As pronouns require local inferences, readers may make regressions to allow for additional processing time following a pronoun. It is a partial limitation of this study that we cannot fully disentangle wrap-up effects from effects that are purely related to anaphor processing. Future studies in this direction may want to use sentences with additional linguistic material between the pronoun and the sentence-final region to enable a better distinction between anaphor processing and end-of-sentence effects.

Taken together, we can conclude from the results of our study that adults' and children's processing of pronouns and repeated names is more similar than expected. We replicated the RNP using eye tracking with adults and saw generally similar anaphor type effects in 9-year-old children. The finding of an RNP in children suggests sensitivity to discourse-level information during online reading processing in beginning readers.

Acknowledgements

We wish to thank Sophia Tischer for her assistance in data collection and preprocessing.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review*, *106*, 748–765.
- Almor, A., & Nair, V. A. (2007). The form of referential expressions in discourse. *Language and Linguistics Compass*, *1*, 84–99.
- Ariel, M. (2001). Accessibility theory: An overview. *Text Representation: Linguistic and Psycholinguistic Aspects*, *8*, 29–87.
- Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using Eigen and R syntax (2011). R package version 0.999375-42. Retrieved from <https://www.scienceopen.com/document?vid=56411f33-a78f-498a-8f6d-876f92bd01cf>
- Blythe, H. I., & Joseph, H. (2011). Children's eye movements during reading. In S. Liversedge, I. Gilchrist & S. Everling (Eds.), *Oxford handbook on eye movements* (pp. 643–662). Oxford, UK: Oxford University Press.
- Drieghe, D., Desmet, T., & Brysbaert, M. (2007). How important are linguistic factors in word skipping during reading? *British Journal of Psychology*, *98*, 157–171.
- Drieghe, D., Brysbaert, M., Desmet, T., & De Baecke, C. (2004). Word skipping in reading: On the interplay of linguistic and visual factors. *European Journal of Cognitive Psychology*, *16*, 79–103.
- Ehrlich, M. F., Remond, M., & Tardieu, H. (1999). Processing of anaphoric devices in young skilled and less skilled comprehenders: Differences in metacognitive monitoring. *Reading and Writing*, *11*, 29–63.
- Engelen, J. A., Bouwmeester, S., de Bruin, A. B., & Zwaan, R. A. (2014). Eye movements reveal differences in children's referential processing during narrative comprehension. *Journal of Experimental Child Psychology*, *118*, 57–77.
- Fox, J., Friendly, M., & Weisberg, S. (2013). Hypothesis tests for multivariate linear models using the car package. *The R Journal*, *5*(1), 39–52.
- Fukumura, K., & van Gompel, R. P. (2015). Effects of order of mention and grammatical role on anaphor resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 501–525.
- Garnham, A., Oakhill, J., & Johnson-Laird, P. N. (1982). Referential continuity and the coherence of discourse. *Cognition*, *11*, 29–46.
- Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition*, *32*, 99–156.
- Givón, T. (1983). *Topic continuity in discourse*. Amsterdam, The Netherlands: John Benjamins.
- Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, *17*, 311–347.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics III: Speech acts* (pp. 41–58). New York, NY: Academic Press.
- Grosz, B. J., Joshi, A. K., & Weinstein, S. (1983, June 15-17). *Providing a unified account of definite noun phrases in discourse*. Proceedings of the 21st Annual Meeting on Association for Computational Linguistics, Cambridge, MA.

- Grosz, B. J., Weinstein, S., & Joshi, A. K. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21, 203–225.
- Hothorn, T., Bretz, F., Westfall, P., Heiberger, R. M., Schuetzenmeister, A., Scheibe, S., & Hothorn, M. T. (2015). Package “multcomp.” Version 1.4-0. Retrieved from <https://cran.r-project.org/web/packages/multcomp/multcomp.pdf>
- Joseph, H. S., Bremner, G., Liversedge, S. P., & Nation, K. (2015). Working memory, reading ability and the effects of distance and typicality on anaphor resolution in children. *Journal of Cognitive Psychology*, 27, 622–639.
- Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.
- Kennison, S. M., & Gordon, P. C. (1997). Comprehending referential expressions during reading: Evidence from eye tracking. *Discourse Processes*, 24, 229–252.
- Lenhard, W., & Schneider, W. (2006). *ELFE 1-6: ein Leseverständnistest für Erst- bis Sechstklässler* [ELFE 1-6: A reading comprehension test for first through sixth grade]. Göttingen, Germany: Hogrefe.
- Lezama, C. G. (2015). A review of the repeated name penalty: Implications for null subject languages. *Revista Linguistica*, 8(2), 22–34.
- Megherbi, H., & Ehrlich, M. F. (2009). The on-line interpretation of pronouns and repeated names in seven-year-old children. *Current Psychology Letters: Behaviour, Brain & Cognition*, 25(2). doi:cpl.revues.org/4895
- Moll, K., & Landerl, K. (2010). *SLRT-II: Lese- und Rechtschreibtest; Weiterentwicklung des Salzburger Lese- und Rechtschreibtests (SLRT)* [SLRT II: reading and spelling test; advancement of the Salzburg reading and spelling test (SLRT)]. Göttingen, Germany: Hogrefe.
- Oakhill, J., & Yuill, N. (1986). Pronoun resolution in skilled and less-skilled comprehenders: Effects of memory load and inferential complexity. *Language and Speech*, 29, 25–37.
- Peters, S. A., Boiteau, T. W., & Almor, A. (2016). Semantic relations cause interference in spoken language comprehension when using repeated definite references, not pronouns. *Frontiers in Psychology*, 7, 214.
- Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *Quarterly Journal of Experimental Psychology* 53(4), 1061–1080.
- R Development Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Schimke, S. (2015). Die rezeptive Verarbeitung von Markierungen der Diskurskohärenz bei Grundschulkindern mit Deutsch als Erst- oder Zweitsprache [The processing of markers of discourse coherence of primary school children with German as their first or second language]. In G. Pagonis & H. Klages (Eds.), *Linguistisch fundierte Sprachförderung und Sprachdidaktik: Grundlagen, Konzepte, Desiderate* (pp. 303–324). Berlin; Boston: De Gruyter.
- Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A., & Kliegl, R. (2015). childLex: A lexical database of German read by children. *Behavior Research Methods*, 47, 1085–1094.
- Shapiro, A., & Milkes, A. (2004). Skilled readers make better use of anaphora: A study of the repeated-name penalty on text comprehension. *Electronic Journal of Research in Educational Psychology*, 2, 161–180.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York, NY: Academic Press.
- Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, 6, 292–297.
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123, 162–185.