

Ina Lehmkuhle* and Sarah Schimke

Are preschool children sensitive to the function of accessibility markers? A visual world study with German-speaking three- to four-year-olds

<https://doi.org/10.1515/lingvan-2022-0066>

Received May 24, 2022; accepted December 20, 2022; published online June 5, 2023

Abstract: Little is known about when children understand the function of anaphoric referring expressions to signal different degrees of accessibility of discourse referents. This visual world study investigates German-speaking three- to four-year-olds' online processing and offline interpretation of repeated names and personal pronouns in a context where reference is made to highly accessible discourse referents. Repeated names are markers of low accessibility, whereas personal pronouns are preferentially used to refer to highly accessible referents. For online processing, results showed a significant effect of referring expression: children looked at the target picture more often after hearing a personal pronoun than after hearing repeated names. Offline results revealed no significant differences between the two conditions. We conclude that German-speaking preschool children are sensitive to the function of accessibility markers during online processing, and suggest that the difference between online and offline results may be due to the different task demands.

Keywords: accessibility; acquisition of reference; German; offline interpretation; online processing; referring expressions

1 Introduction

Referring to referents already mentioned in the discourse is a central means to establish a coherent discourse. It is assumed that the choice of a particular referring expression correlates with the relative degree of discourse referents' accessibility (Ariel 1990): for example, personal pronouns are high accessibility markers, whereas definite noun phrases or repeated names normally refer to referents with low accessibility. As in English, repeated names in German can be considered as markers of low accessibility. The personal pronoun (*er*) is considered a marker for high accessibility in German. It represents the most reduced form in most contexts, as topic drop is only allowed in very restricted contexts (Fries 1988). Furthermore, it contrasts with a second more marked personal pronoun (*der*), as well as with demonstrative pronouns (*dieser*, *jener*), which can all be considered to mark lower degrees of accessibility than *er* (Ellert 2010). There are several factors such as grammatical function and order of mention that make a discourse referent more or less accessible (for an overview, see Arnold 2010). For example, when resolving ambiguous pronouns, adults prefer subject antecedents over object antecedents, and first-mentioned antecedents over second-mentioned antecedents (Järvikivi et al. 2005), suggesting that subject and first-mentioned referents have a higher degree of accessibility than non-subject and last-mentioned referents.

Sarah Schimke is now working at Ludwig-Maximilians-Universität in Munich.

***Corresponding author: Ina Lehmkuhle**, Institut für Germanistik, Universität Osnabrück, Neuer Graben 40, Osnabrück, 49069, Germany, E-mail: ina.lehmkuhle@uos.de. <https://orcid.org/0000-0003-0028-7972>

Sarah Schimke, Institut für Sprache, Literatur und Kultur, TU Dortmund, Dortmund, Nordrhein-Westfalen, Germany

There are several studies concerned with children's development of processing referring expressions. Most of these studies investigate whether children are sensitive to the relative degree of accessibility of discourse referents when processing ambiguous personal pronouns by manipulating various accessibility factors (Arnold et al. 2007; Hartshorne et al. 2015; Järvikivi et al. 2014; Klages and Gerwien 2014; Pyykkönen et al. 2010; Song and Fisher 2005, 2007). These studies indicate that children are sensitive to the relative degree of accessibility of discourse referents from an early age. For example, Pyykkönen et al. (2010) showed that English-speaking three-year-old children are sensitive to syntactic and semantic accessibility factors when processing ambiguous personal pronouns in a visual world study with two potential antecedents. The children looked at pictures of both referents more often in a condition with high-transitive verbs such as *beat* than in a condition with low-transitive verbs such as *see*. Pyykkönen et al. (2010) assumed that this pattern is due to the referents of high-transitive verbs having more prototypical agent and patient properties, which are presumed to increase the accessibility of both referents. Moreover, there were generally more looks to subject antecedents than object antecedents in both conditions. Similarly, Hartshorne et al. (2015) investigated whether children show a first-mention bias when processing ambiguous personal pronouns. Like an adult control group, English-speaking five-year-old children looked more frequently at the picture of the first-mentioned antecedent than that of the last-mentioned one when listening to ambiguous pronouns with two potential antecedents.

While there is much literature on the online processing of ambiguous personal pronouns, there are only a few studies which have examined more directly when children know that a core function of referring expressions is to signal different degrees of accessibility. One such study, by Skarabela and Ota (2017), manipulated the referring expression while maintaining the degree of accessibility. Using the cross-modal preferential-looking paradigm, they investigated how English-speaking children aged one and a half to two years processed definite noun phrases and personal pronouns in comparison to indefinite noun phrases. The expressions referred to a previously introduced, highly accessible referent in sentences like: *Look, a ball. Can you see a hat/the ball/it?* The results showed that in the definite noun phrase or pronoun conditions, the two-year-old children looked more often to the given referents (e.g., *ball*) than to the unfamiliar referents (e.g., *hat*), similar to an adult control group. The younger children (1.5–1.7 years) behaved differently: they showed a preference for the correct picture only for the definite noun phrase condition, but not the pronoun condition. The authors conclude that children understand that personal pronouns refer to highly accessible referents at some point between one and a half and two years of age.

For German, no studies exist on this topic for such young children. There are only reading studies with older children, which have concluded that children recognize the function of various accessibility markers in discourse during online processing, based on the fact that they showed a repeated-name penalty (Gordon et al. 1993) similar to adults. Gordon et al. (1993) showed in a series of reading experiments that repeated names lead to increased reading times compared to personal pronouns when they refer to highly accessible referents, as they do not agree with the degree of accessibility of the referents. In a self-paced reading experiment, Schimke (2014) found a repeated-name penalty in German-speaking ten-year-olds: repeated names were read more slowly than pronouns. Similarly, Eilers et al. (2019) examined monolingual German-speaking children aged nine and ten years and an adult control group in an eye-tracking during reading experiment. They found a repeated-name penalty for both adults and children. Participants' gaze duration increased in the region after the anaphora for repeated names compared to personal pronouns, indicating increased difficulties integrating the repeated names. Children and adults also showed more regressions to previous regions of the text in the repeated names condition, presumably because they were trying to repair the unexpected occurrence of the repeated name. The two studies suggest that German-speaking children at the age of nine and ten recognize the function of different referring expressions in discourse. However, given that there are no studies on younger children, it remains open at what age this capacity develops. There is some evidence that the development of German-speaking children may differ from that of English-speaking children with respect to the acquisition of the functions of referring expressions. While English is neither a pro-drop nor a zero-topic language, German is a zero-topic language (Huang 1984), in which the topic can be dropped in certain cases (Fries 1988). For example, a study by Hickmann and Hendriks (1999) found that English-speaking adults and four- to ten-year-olds used more personal pronouns than zero pronouns to maintain animated – highly accessible – characters in elicited narratives. German-speaking adults

and ten-year-olds, on the other hand, used more zero pronouns than personal pronouns in this context. The younger age group (four to seven years old) behaved differently: like the English-speaking participants, they used more personal pronouns than zero pronouns. Moreover, as mentioned above, German uses two series of personal pronouns (*er* and *der*; Ellert 2010), which differ in the degree of accessibility that they signal. Given these greater complexities of the German system, it is conceivable that German-speaking children need more time to learn the different functions of the different pronoun types than English-speaking children.

Another aspect that needs to be considered when studying children is the method and the tasks used in the studies. Often, a dissociation between online processing and offline interpretation is observed in pronoun interpretation (for an overview, see Sekerina 2015): when using offline methods such as sentence-picture matching tasks which capture the final interpretation of pronouns, children were found to behave non-target-like at the age of six (Sekerina 2015), whereas the online studies cited above showed more target-like processing. The resulting dissociation was found, for instance, by Sekerina et al. (2004). They examined how English-speaking four- and seven-year olds and an adult control group process (ambiguous) reflexive and personal pronouns in an eye-tracking experiment (online data) which included a naming task (offline data). Unlike the adults, the children interpreted both the personal and reflexive pronouns almost exclusively as referring to a sentence-internal antecedent. Online, however, they showed a similar pattern as the adults: they looked at the within-sentence referent significantly more often in the condition with the reflexive pronouns than in the condition with the personal pronouns. The online method was thus able to make linguistic knowledge visible that was not revealed in the offline task. For this reason, language acquisition studies should include both online and offline measurements, as offline data alone can often lead to an underestimation of children's abilities (Sekerina 2015).

2 This study

The present study investigates three- to four-year-old German-speaking children's online processing and offline interpretation of repeated names and personal pronouns referring to highly accessible discourse referents.¹ We collect both online and offline data. We use the visual world paradigm, which goes back to Tanenhaus et al. (1995) and allows for the observation of the processing of the referring expressions in real time, as well as a sentence-picture matching task embedded in it, which captures the children's final interpretation. The visual world paradigm is particularly well suited for studying children because it does not demand any special skills from them. For example, unlike the self-paced-reading experiments in the repeated-name penalty context, no reading skills are required. Instead, the children only have to listen and look at the pictures, which is a relatively natural behaviour. Our research questions are as follows:

- Q1 Are three- to four-year-old German-speaking children sensitive to the functional difference between personal pronouns and repeated names during online processing when they refer to highly accessible referents?
- Q2 How does online processing differ from offline interpretation?

If the children are sensitive to the difference between personal pronouns and repeated names, as the English-speaking children in Skarabela and Ota's (2017) study were, they should prefer personal pronouns over repeated names because they match the accessibility of the discourse referents who are highly accessible. We assume that this should be reflected in a greater number of fixations towards the target picture representing the antecedents of the anaphora after hearing a personal pronoun than after hearing the repeated names. However, if the children have not yet learnt the function of personal pronouns to signal a high degree of accessibility, it could be the case that there will be no difference between the two conditions.

As for the difference between online processing and offline interpretations, it is conceivable that the online data will show more linguistic knowledge than the offline data. In the offline task, children may not manage to

¹ This paper is based on one chapter of the first author's doctoral dissertation (Lehmkuhle 2022).

reliably access the interpretation they have built up during online processing, possibly due to their restricted working memory capacities, as was assumed by Sekerina et al. (2004).

3 Methods

3.1 Participants

Participants in the experiment were 27 German-speaking children (17 female, 10 male; age range = 3;1–4;10 years; mean age = 3;10 years) from kindergartens in a rural area in northern Germany.² Prior to the study, parents gave their written consent and provided background information by completing a questionnaire. All children were monolingual speakers of German and had normally developed hearing and vision, with the exception of two children whose visual impairment was corrected with glasses. In addition to the experiment, the children took a grammar comprehension test (TROG-D) developed by Fox-Boyer et al. (2016) showing that, on average, the children had an understanding of grammar appropriate to their age (range = 28–61; mean: 46.48).³

3.2 Materials

Materials consisted of a zoo story that was presented to the participants auditorily: the protagonists Mimi and Benny spend a day at the zoo on the occasion of Mimi's birthday; Benny is familiar with the zoo and introduces Mimi to the different animals by name on a tour around the zoo; on their way, they also meet various people such as school classmates or the zoo staff. Ten test items and five filler items are embedded in the story.⁴ A test

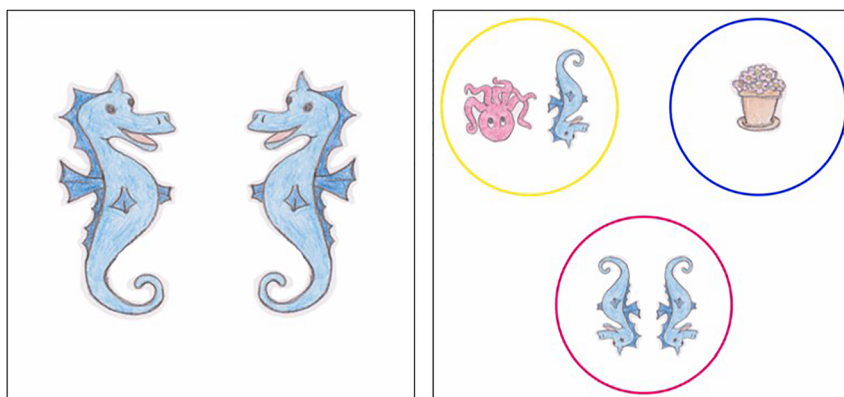


Figure 1: Example test item: screen context sentence (*left*) and screen critical sentence (*right*).

² The sample sizes of comparable previous studies working with the visual world paradigm are: 24, 23, and 16 in the studies reported in Arnold et al. (2007); 22, 24, and 20 in Song and Fisher (2005); 19 in Pyykkönen et al. (2010); 39 in Järvikivi et al. (2014); 45 in Hartshorne et al. (2015); 18, 23, and 18 in Klages and Gerwien (2014); and 16 and 16 in Skarabela and Ota (2017). In all these studies, effects were found. They are not directly comparable to our study because they investigated ambiguous pronouns in different contexts (e.g., different degrees of accessibility in terms of order of mention). Since our study deals with a rather basic question in comparison to these studies (high degree of accessibility in all context sentences by changing only the referring expression), we considered a sample size lying within this range as appropriate.

³ The test was standardized on 870 monolingual children growing up with German as their first language, aged between 3;0 and 10;11 years. The average T-score in the test is 50, with a standard deviation of 10 points.

⁴ A list with the test and the filler items as well as the images used can be found in the appendix of the first author's doctoral dissertation (Lehmkuhle 2022).

item (see Figure 1) consists of a context sentence (1), which introduces two referents, as well as a critical sentence manipulated with respect to the referring expression in the conditions “repeated names” (2a) and “pronoun” (2b), in which these same referents are referred to again:

- (1) *Das sind Sarah und Susi.*
 ‘These are Sarah and Susi.’
- (2) a. *Sarah und Susi machen einen Kopfstand.*
 ‘Sarah und Susi do a headstand.’
 b. *Sie machen einen Kopfstand.*
 ‘They do a headstand.’

The auditory presentation of the context sentence is preceded by a picture showing the introduced referents, and the presentation of the critical sentence is followed by an image showing three pictures in coloured circles.

The critical image consists of a target, a competitor, and a distractor. In a randomized order, these appear equally often in each of the three positions. The three circles have the same distance from each other and are of the same size. The target consists of the two animals previously presented in the context sentence performing an action together, the competitor consists of one of the previously introduced animals and another animal performing this action together, and the distractor contains an inanimate object.

3.3 Procedure

The experiment took place in a quiet room in the child’s kindergarten. Each child was tested individually. A Tobii TX300 eye tracker recorded the eye movements of the participants. The stimuli used had been previously created using the video editing software Windows Movie Maker as WMV files with a resolution of $1,920 \times 1,080$ pixels and were presented on an external screen with a resolution of $1,920 \times 1,080$ pixels. The Tobii Pro Studio software, which ran in parallel with the collection on a Dell laptop, recorded the data. Sound was played by an external Bluetooth speaker. The participants sat in front of the external screen and the experimenter to their left in front of the laptop. The children were told that they would hear a magic story with pictures in coloured circles appearing on the screen at some points. They were also told that the story would stop and only continue when they named the colour of the circle that best matched what they had just heard.⁵ The experimenter noted the children’s responses during the study. The auditory stimulus appeared 2,600 ms after the presentation of the critical image. This value was chosen following a study by Cristante (2016), where the same preview time was used for a visual world study with children. This preview time gives participants the chance to look at the pictures and build up a mental representation of the visual context before they listen to the verbal stimulus. In three warm-up items, the three colours of the circles were practised in advance. Children who did not respond at the critical point were asked the supportive question: *Welche Farbe passt am besten?* ‘Which colour fits best?’. After the experiment, the children completed the previously mentioned grammar comprehension test, TROG-D. In total, the session lasted about 20 min.

4 Results

In order to evaluate the explicit interpretation of the referring expressions, the answers of the children were analysed with respect to the target picture (offline data). The analysis of the eye movements (online data) was done with the freely available R package eyetrackingR (Dink and Ferguson 2015).

⁵ Since children of this age usually have difficulty sitting still for a while without moving – which is obviously necessary for tracking eye movements – we decided to have the children name the colour of the circle. We were concerned that pointing to the pictures would cause movement, possibly leading to trackloss.

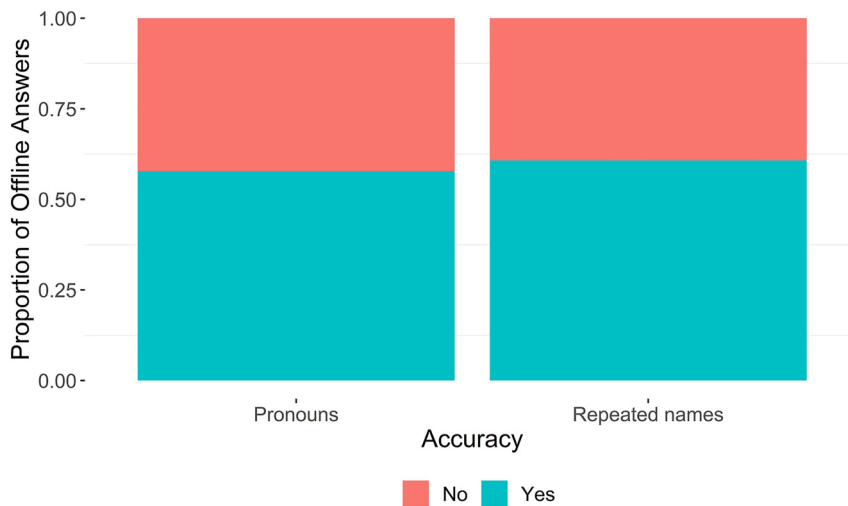


Figure 2: Proportion of accurate and inaccurate offline choices for pronouns and repeated names.

4.1 Offline results

In total, the children produced 270 answers in the sentence-picture matching task. After hearing a pronoun, they selected the target picture in 58 percent ($n = 78$) of the cases, and after hearing the repeated names in 61 percent ($n = 82$) of the cases (see Figure 2).

Note that there was huge variation within the group. In order to find out whether this performance was above chance and to what extent the referring expression influenced the accuracy of the answers, we performed a statistical analysis with accuracy as the dependent variable, and participant and item as random factors. In addition to having referring expression as an independent variable, we included the results of the TROG-D test (TROG-D) and the age of the children in months (age) as independent variables to evaluate the extent to which these factors could possibly explain the large differences between the children. Since these two variables are continuous variables, we normalized them in advance. We also included interaction effects between each of these two variables and referring expression. We built a linear mixed-effects model, using the R package *buildmer* which finds the maximal model that still converges by simplifying the random effects structure using stepwise elimination when the full maximal model does not converge (Voeten 2021). This package supports models which can be fitted by (g)lm and (g)lmer, among others (R package *lme4*; Bates et al. 2015). We only report the final model. The results (see Table 1) show that there was a significant preference for the target picture compared to the

Table 1: Model summary for the linear mixed-effects model: accuracy offline choices.

Random effects	Variance	SD	Correlation	
Participant (intercept)	0.073	0.270		
Item (intercept)	0.007	0.082		
TROG-D	0	0.020	-0.530	
Fixed effects	Estimate	SE	<i>t</i>	<i>p</i>
Intercept	0.607	0.072	8.388	0***
Referring expression (pronoun vs. repeated names)	-0.030	0.071	-0.415	0.687
TROG-D	0.044	0.063	0.701	0.489
Age	0.047	0.062	0.756	0.454
Referring expression × TROG-D	-0.054	0.050	-1.120	0.291
Referring expression × age	-0.032	0.049	-0.656	0.513

Notes. Linear mixed-effects model with random intercept for participant. Model fit by maximum likelihood. All values have been rounded to three decimal places. *** $p < 0.001$.

incorrect choices (e.g., distractor or competitor). Furthermore, the children performed above chance in both conditions, which is reflected in the significant intercept. However, none of the three experimental factors had an effect on the accuracy of the answers. Thus, in the offline interpretation of the referring expressions, we see no significant difference between pronouns and repeated names. Moreover, neither the age of the children nor the results from the TROG-D can explain the variation in accuracy.

4.2 Online results

To analyse the online processing of the referring expression, we conducted three statistical analyses on the target fixation between 500 and 2000 ms after the onset of the anaphora,⁶ using target fixation as the dependent variable, referring expression as the independent variable, and participant and item as random effects. Looking time was averaged across the time window. Similar to the analysis of the offline data, we used the R package *buildmer* to calculate the maximal model. We only report the final models.

Table 2: Model summary for the linear mixed-effects model: target fixation between 500 and 2000 ms after anaphora onset. All trials.

Random effects	Variance	SD		
Participant (intercept)	1.430	1.196		
Fixed effects	Estimate	SE	t	p
Intercept	−1.131	0.825	−1.370	0.177
Referring expression (pronoun vs. repeated names)	2.126	1.044	2.037	0.045*

Notes. Linear mixed-effects model with random intercept for participant. Model fit by maximum likelihood. All values have been rounded to three decimal places. * $p < 0.05$.

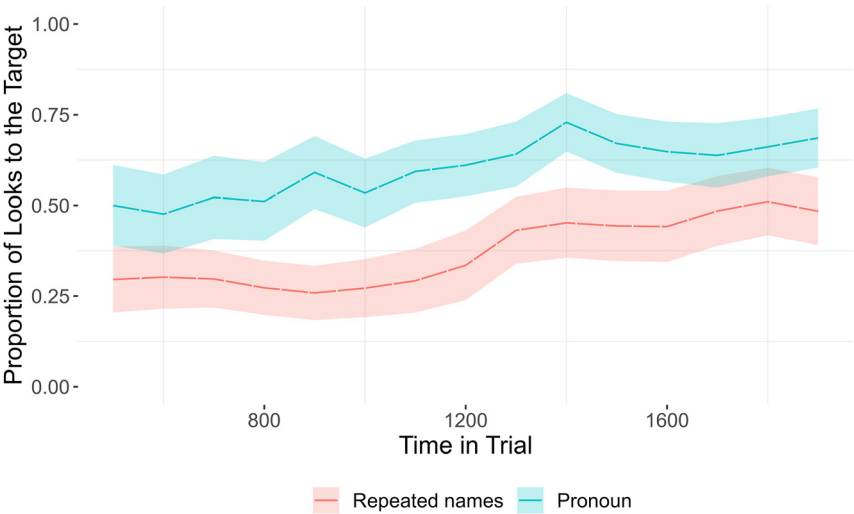


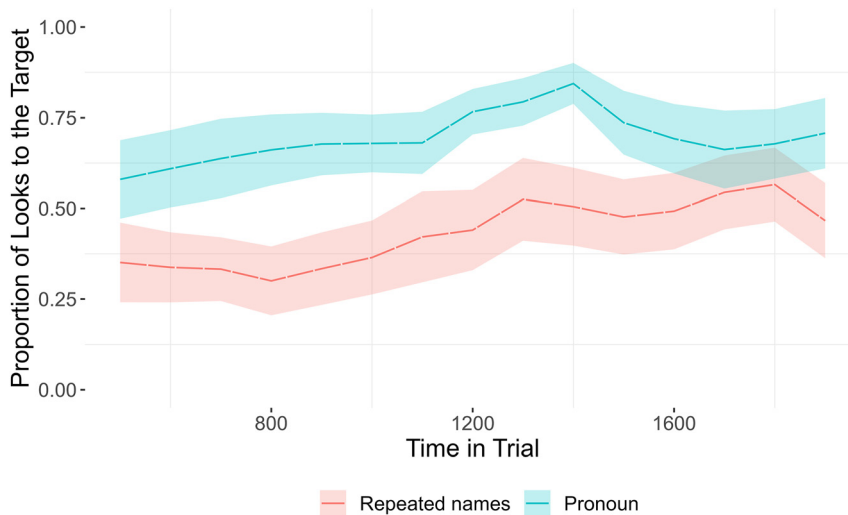
Figure 3: Proportion of looks to the target between 500 and 2000 ms after the onset of the anaphora for pronouns and repeated names (all trials).

⁶ We based our choice of time window on previous studies. For example, Järvikivi et al. (2014) investigated German-speaking children of the same age as the children in our study and chose 200–2000 ms after pronoun onset as the time window. However, since we investigated not only pronouns but repeated names as well, we decided to take 500 ms after anaphora onset as the starting point, as repeated names need more time to unfold than pronouns.

Table 3: Model summary for the general linear model: target fixation between 500 and 2000 ms after anaphora onset. Only accurate offline choices.

Coefficients	Estimate	SE	<i>t</i>	<i>p</i>
Intercept	−0.798	0.960	−0.831	0.410
Referring expression (pronoun vs. repeated names)	2.889	1.323	2.184	0.033*

Notes. General linear model. All values have been rounded to three decimal places. * $p < 0.05$.

**Figure 4:** Proportion of looks to the target between 500 and 2000 ms after the onset of the anaphora for pronouns and repeated names (only trials with accurate offline choices).

The first model includes *all trials*, regardless of which answer the children gave offline. The results (see Table 2) show a significant effect for referring expression ($p < 0.05$). As can be seen in Figure 3, the children looked at the target picture significantly more often after the presentation of a pronoun than after the presentation of the repeated names.

The second model only includes *trials associated with accurate offline choices*. Again, the analysis (see Table 3) shows a significant effect for referring expression ($p < 0.05$). Figure 4 illustrates that the children looked at the target significantly more often after hearing a pronoun than after hearing the repeated names, also in those cases where they ultimately chose the target picture for both expressions.

Finally, the third model only includes *trials associated with inaccurate offline choices*. The analysis (see Table 4) shows a marginally significant effect of referring expression ($p = 0.05$). As can be seen in Figure 5, even when the children selected an incorrect picture, they looked marginally significantly more frequently at the target after the presentation of a pronoun than after the presentation of the repeated names.

Table 4: Model summary for the linear mixed-effects model: target fixation between 500 and 2000 ms after anaphora onset. Only inaccurate offline choices.

Random effects	Variance	SD		
Participant (intercept)	0.869	0.932		
Fixed effects	Estimate	SE	<i>t</i>	<i>p</i>
Intercept	−2.467	1.263	−1.953	0.067
Referring expression (pronoun vs. repeated names)	3.964	1.911	2.075	0.053

Notes. Linear mixed-effects model with random intercept for participant. Model fit by maximum likelihood. All values have been rounded to three decimal places.

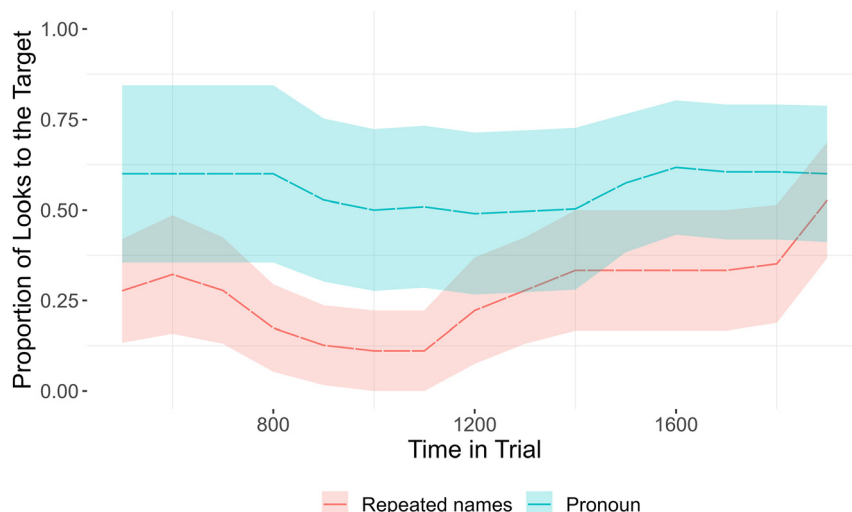


Figure 5: Proportion of looks to the target between 500 and 2000 ms after the onset of the anaphora for pronouns and repeated names (only trials with inaccurate offline choices).

5 Discussion and conclusion

The present study investigated three- to four-year-old German-speaking children's online processing and offline interpretation of repeated names and personal pronouns referring to highly accessible discourse referents. We predicted that the children would prefer personal pronouns over repeated names during online processing if they were sensitive to the difference between these two referring expressions, and that this would be visible in an increased number of fixations towards the target after hearing a personal pronoun compared to after hearing the repeated names. In addition, we expected a dissociation between online processing and offline interpretation, with online data revealing more linguistic knowledge than offline data. Both predictions were borne out.

The online results show that the children were sensitive to the different functions of pronouns and repeated names. They looked at the target significantly more often after hearing a personal pronoun which matched the degree of accessibility of the discourse referents than after hearing the repeated names. We interpret this as evidence that they are sensitive to the fact that personal pronouns are more appropriate than repeated names for referring to highly accessible discourse referents. This is the first study to show this for German-speaking children at this young age. As expected, we observed a dissociation between online processing and offline interpretation. Children's offline choices were above chance, slightly favouring repeated names (61 percent) over personal pronouns (58 percent), with no significant difference between the two conditions. There was large individual variation that could neither be attributed to the children's age nor to the results of the TROG-D test.

The results are in line with previous research. On the one hand, the findings confirm that children are already sensitive at an early age to the fact that pronouns refer to highly accessible discourse referents (Arnold et al. 2007; Hartshorne et al. 2015; Järviö et al. 2014; Klages and Gerwien 2014; Pyykkönen et al. 2010; Song and Fisher 2005, 2007) and are more appropriate for this purpose than other referring expressions such as repeated names (Eilers et al. 2019; Schimke 2014; Skarabela and Ota 2017). On the other hand, the results illustrate that it is important to complement offline measures with online measures to investigate preschool children's perceptual abilities with regard to understanding referring expressions.

It has already been observed that children's performance is strongly influenced by the demands of offline methods (Sekerina 2015). That this also seems to be the case here is shown by the fact that children prefer pronouns in online processing, even in those cases where they ultimately chose a non-target picture. One reason for this dissociation could be that children's executive functions, which include working memory, for instance, are not yet as fully developed as those of adults (Hopp and Schimke 2018): reduced working memory capacities could lead to a child selecting an image that does not match the content of an associated sentence, even if they have built up a correct interpretation when processing the sentence. In future studies, it would therefore be reasonable to take into account individual differences with regard to executive functions such as the participants'

working memory capacity. Note that the dissociation between our online and offline data might have been exacerbated by the particular difficulty of our offline task. After all, not only did children have to find the correct picture, but they also had to name the colour of the circle that enclosed it. Note, however, that the overall accuracy was above chance. While the task was certainly more complex than a pointing task, for instance, we thus do not think that it masked all preferences altogether. Moreover, as a dissociation between the knowledge observable in online and offline tasks has been observed by others as well when using simpler offline tasks (e.g., Sekerina 2015), we also do not think that the difference between our online and offline results is exclusively due to the particular complexity of the offline task. Nevertheless, the dissociation is reason to caution against making generalizations based on the results obtained with one method only, and a reminder of the benefits of combining different methods in language acquisition studies.

In conclusion, this study shows that monolingual German-speaking children aged three to four years notice the functional difference between personal pronouns and repeated names in online processing. They preferred personal pronouns over repeated names when these referred to highly accessible discourse referents, showing that they are sensitive to the fact that referring expressions signal the relative degree of accessibility of discourse referents. This was shown regardless of whether they selected the correct picture or not.

References

- Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge.
- Arnold, Jennifer E. 2010. How speakers refer: The role of accessibility. *Language and Linguistics Compass* 4(4). 187–203.
- Arnold, Jennifer E., Sarah Brown-Schmidt & John Trueswell. 2007. Children's use of gender and order-of-mention during pronoun comprehension. *Language and Cognitive Processes* 22(4). 527–565.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.
- Cristante, Valentina. 2016. *The processing of non-canonical sentences in children with German as a first or second language and German adults*. Münster: Westfälische Wilhelms-Universität Münster Dissertation.
- Dink, Jacob & Brock Ferguson. 2015. eyetrackingR: An R library for eye-tracking data analysis. <http://www.eyetracking-r.com> (accessed 6 July 2021).
- Eilers, Sarah, Simon P. Tiffin-Richards & Sascha Schroeder. 2019. The repeated name penalty effect in children's natural reading: Evidence from eye tracking. *Quarterly Journal of Experimental Psychology* 72(3). 403–412.
- Ellert, Miriam. 2010. *Ambiguous pronoun resolution in L1 and L2 German and Dutch* (MPI Series in Psycholinguistics 58). Wageningen: Ponsen & Looijen.
- Fox-Boyer, Annette, Thomas Bäumer, Madeleine Müller & Silke Merzbecher (eds.). 2016. *TROG-D: Test zur Überprüfung des Grammatikverständnisses*, 7th edn. Idstein: Schulz-Kirchner Verlag.
- Fries, Norbert. 1988. Über das Nulltopik im Deutschen. *Sprache und Pragmatik* 3. 15–49.
- Gordon, Peter C., Barbara J. Grosz & Laura A. Gilliom. 1993. Pronouns, names, and the centering of attention in discourse. *Cognitive Science* 17(3). 311–347.
- Hartshorne, Joshua K., Rebecca Nappa & Jesse Snedeker. 2015. Development of the first-mention bias. *Journal of Child Language* 42(2). 423–446.
- Hickmann, Maya & Henriëtte Hendriks. 1999. Cohesion and anaphora in children's narratives: A comparison of English, French, German, and Mandarin Chinese. *Journal of Child Language* 26(2). 419–452.
- Hopp, Holger & Sarah Schimke. 2018. Einleitung: Fragestellungen, Modelle und Methoden in der L2-Verarbeitungsforschung. In Sarah Schimke & Holger Hopp (eds.), *Sprachverarbeitung im Zweitspracherwerb*, 1–28. Berlin & Boston: De Gruyter Mouton.
- Huang, C. T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15(4). 531–574.
- Järvikivi, Juhani, Roger P. G. van Gompel, Jukka Hyönä & Raymond Bertram. 2005. Ambiguous pronoun resolution: Contrasting the first-mention and subject-preference accounts. *Psychological Science* 16(4). 260–264.
- Järvikivi, Juhani, Pirta Pyykkönen-Klauck, Sarah Schimke, Saveria Colonna & Barbara Hemforth. 2014. Information structure cues for 4-year-olds and adults: Tracking eye movements to visually presented anaphoric referents. *Language, Cognition and Neuroscience* 29(7). 877–892.
- Klages, Hana & Johannes Gerwien. 2014. Verstehen anaphorischer Personalpronomina im DaZ- und DaM-Erwerb. In Hana Klages & Giulio Pagonis (eds.), *Linguistisch fundierte Sprachförderung und Sprachdidaktik: Grundlagen, Konzepte, Desiderate*, 71–98. Berlin: De Gruyter Mouton.
- Lehmkuhle, Ina. 2022. *Referenzielle Kohärenz im Erstspracherwerb: Untersuchungen zur Verarbeitung und Produktion anaphorischer Referenz*. Osnabrück: University of Osnabrück Dissertation.

- Pyykkönen, Pirita, Danielle Matthews & Juhani Järvikivi. 2010. Three-year-olds are sensitive to semantic prominence during online language comprehension: A visual world study of pronoun resolution. *Language and Cognitive Processes* 25(1). 115–129.
- Schimke, Sarah. 2014. Die rezeptive Verarbeitung von Markierungen der Diskurskohärenz bei Grundschulkindern mit Deutsch als Erst- oder Zweitsprache. In Hana Klages & Giulio Pagonis (eds.), *Linguistisch fundierte Sprachförderung und Sprachdidaktik: Grundlagen, Konzepte, Desiderate*, 193–213. Berlin: De Gruyter Mouton.
- Sekerina, Irina A. 2015. Online evidence for children's interpretation of personal pronouns. In Ludovica Serratrice & Shanley E. M. Allen (eds.), *The acquisition of reference*, 213–239. Amsterdam: John Benjamins.
- Sekerina, Irina A., Karin Stromswold & Arild Hestvik. 2004. How do adults and children process referentially ambiguous pronouns? *Journal of Child Language* 31(1). 123–152.
- Skarabela, Barbora & Mitsuhiro Ota. 2017. Two-year-olds but not younger children comprehend it in ambiguous contexts: Evidence from preferential looking. *Journal of Child Language* 44(1). 255–268.
- Song, Hyun-joo & Cynthia Fisher. 2005. Who's "she"? Discourse prominence influences preschoolers' comprehension of pronouns. *Journal of Memory and Language* 52(1). 29–57.
- Song, Hyun-joo & Cynthia Fisher. 2007. Discourse prominence effects on 2.5-year-old children's interpretation of pronouns. *Lingua* 117(11). 1959–1987.
- Tanenhaus, Michael K., Michael Spivey-Knowlton, Kathleen M. Eberhard & Julie C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268(5217). 1632–1634.
- Voeten, Cesko C. 2021. Using "buildmer" to automatically find & compare maximal (mixed) models. 19 August 2021. <https://cran.r-project.org/web/packages/buildmer/vignettes/buildmer.html> (accessed 23 May 2022).