Journal of Experimental Child Psychology 235 (2023) 105690

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Even young children make multiple predictions in the complex visual world

Linda Sommerfeld [a,*], Maria Staudte [b,1], Nivedita Mani [c], Jutta Kray [a]

[a] Department of Psychology, Saarland University, 66123 Saarbrücken, Germany
[b] Department of Computational Linguistics and Phonetics, Saarland University, 66123 Saarbrücken, Germany
[c] Georg Elias Müller Institute for Psychology, University of Göttingen, 37073 Göttingen, Germany

ARTICLE INFO

Article history:
Received 8 June 2022
Revised 14 February 2023

Keywords:
Prediction
Visual world
Children
Language comprehension
Eye-tracking

ABSTRACT

Children can anticipate upcoming input in sentences with semantically constraining verbs. In the visual world, the sentence context is used to anticipatorily fixate the only object matching potential sentence continuations. Adults can process even multiple visual objects in parallel when predicting language. This study examined whether young children can also maintain multiple prediction options in parallel during language processing. In addition, we aimed at replicating the finding that children's receptive vocabulary size modulates their prediction. German children (5–6 years, n = 26) and adults (19–40 years, n = 37) listened to 32 subject–verb–object sentences with semantically constraining verbs (e.g., "The father *eats* the waffle") while looking at visual scenes of four objects. The number of objects being consistent with the verb constraints (e.g., being *edible*) varied among 0, 1, 3, and 4. A linear mixed effects model on the proportion of target fixations with the effect coded factors condition (i.e., the number of consistent objects), time window, and age group revealed that upon hearing the verb, children and adults anticipatorily fixated the single visual object, or even multiple visual objects, being consistent with the verb constraints, whereas inconsistent objects were fixated less. This provides first evidence that, comparable to adults, young children maintain multiple prediction options in parallel. Moreover, children with larger receptive vocabulary sizes (Peabody Picture Vocabulary Test) anticipatorily fixated potential targets more often than those with smaller ones, showing that

* Corresponding author.
E-mail address: lindas@coli.uni-saarland.de (L. Sommerfeld).

[1] Current address: Bavarian Research Institute for Digital Transformation, 80333 München, Germany.

https://doi.org/10.1016/j.jecp.2023.105690

verbal abilities affect children's prediction in the complex visual world.

Introduction

Language Prediction in Children

Considerable research has shown that, when processing language, adults and children can leverage their semantic knowledge to actively predict upcoming words in the input rather than just passively receiving them (Altmann & Kamide, 1999; Andreu et al., 2013; Borovsky & Creel, 2014; Borovsky et al., 2012; Mani & Huettig, 2012, 2014; Nation et al., 2003; Reuter et al., 2021; see Pickering & Gambi, 2018, for an overview). Thus, these studies found that children from as early as 2 years of age (Mani & Huettig, 2012) as well as older children (Andreu et al., 2013; Mani & Huettig, 2014; Nation et al., 2003) and adults (Ankener et al., 2018; Altmann & Kamide, 1999) can use semantically constraining verbs that allow for only a limited number of arguments to predict as plausible continuations of thus far presented sentences. These studies typically employ the *visual world paradigm*, where participants' eye movements are recorded while the participants look at visual scenes and listen to language (Allopenna et al., 1998; see Huettig et al., 2011, for a review). In developmental research, children are often presented with simple two-object displays while they listen to predictive sentences with semantically constraining verbs (e.g., "The boy *eats* the cake"). Typically, one of the objects presented on-screen (e.g., cake) is consistent with the verb's semantic constraints (e.g., is edible), whereas the other is not (e.g., bird). The finding that children anticipatorily fixate the target object (cake) after hearing the constraining verb (eat) but, importantly, before the target word (cake) is interpreted as evidence of children anticipating upcoming linguistic input (Mani & Huettig, 2012, 2014). In other words, despite their limited word and world knowledge (relative to adults), even young children are able to leverage what they know to anticipate upcoming words and their meanings, that is, the associated semantic and perceptual features (see Pickering & Gambi, 2018, for an overview). This could explain their proficient handling of language, such that the speed and accuracy of language processing derive, at least in part, from the anticipation of upcoming linguistic input based on auditory information (Fernald et al., 2008; Gambi, 2021; Mahr et al., 2015; Mani & Huettig, 2014). Thus, prediction boosts language processing because children can process predictable words even before they are presented, thereby speeding up word recognition for such words.

Moreover, there is ample evidence that also visual information that reflects the meaning of the linguistic input can boost children's and adults' language processing (e.g., Engelhardt et al., 2010; Knoeferle et al., 2005; Scheepers & Crocker, 2004; van Rij, 2012; Weighall & Altmann, 2011; Zhang & Knoeferle, 2012). For instance, it is easier for children (4–5 years) to comprehend rather unfamiliar object–verb–subject sentences (e.g., "The bear is painted by the worm") when they are presented together with visual scenes that depict (e.g., show a bear that is painted by a worm) versus do not depict (e.g., show a bear next to a worm) the meaning of the sentences (Zhang & Knoeferle, 2012). Thus, visual information also seems to affect language comprehension during early childhood.

So far, most studies examining the prediction boost in language processing focus on the semantic relation between the words in the (typically) auditory input while paying less attention to potential contributions of the visual context in which such input is presented. Studies with children mostly present only one visual referent of the predictable word. However, language processing usually takes place in a more complex visual environment (Reuter et al., 2021); thus, other suitable referents from the visual context may also affect what children consider as possible sentence continuations (as has been shown for adults by Ankener et al., 2018). The role of this visual context in our understanding of prediction in young children therefore is still unclear. Against this background, the current study examined the extent to which the visual context in which predictable words are presented influences children's language prediction.

Influence of complex visual contexts

One may assume that only in simple visual scenarios, such as the two-object displays reviewed above, could it be possible for visual information to be processed fast enough to keep pace with prediction (Reuter et al., 2021). However, this is not the case. Prediction can also be observed in more complex visual contexts. When listening to predictive sentences (e.g., "The man *milks* the cow"), adults as well as children aged 3 years and older have been shown to anticipatorily fixate the single visual referent (e.g., cow) that is consistent with the semantic constraints of a verb more than three distractor objects (Andreu et al., 2013; Borovsky et al., 2012; Nation et al., 2003). In a similar study with naturalistic photographs as stimuli, adults and preschoolers have also been shown to anticipatorily fixate the single appropriate visual referent out of more than 15 distractors (Reuter et al., 2021). These findings suggest that adults and even young children efficiently integrate complex visual scenarios into language prediction.

However, in the real world, there may be multiple visual stimuli that are consistent with the constraints of the linguistic input. Here, adults and children may follow a multiple predictions pattern; they might integrate multiple visual cues that are consistent with the input presented thus far in parallel. This would suggest that, when processing language, they maintain multiple prediction options in parallel. Given that correct predictions boost the speed of language processing, maintaining multiple prediction options may allow them to maximally benefit from prediction during language processing (Kuperberg & Jaeger, 2016; van Petten & Luka, 2012). Otherwise, adults and children could follow a one-only prediction pattern and integrate only single visual prediction options into language processing, implying that they avoid possible costs of unfulfilled predictions (Kuperberg & Jaeger, 2016; van Petten & Luka, 2012).

In general, children and adults have already been shown to integrate multiple visual stimuli in parallel into sentence processing. Two pioneering studies of online sentence processing presented children (5 years) and adults with verbal instructions such as "Put the frog [...] in the box" while participants looked at two stuffed frogs and two other objects in front of them (Snedeker & Trueswell, 2004; Trueswell et al., 1999). When hearing the first part of the instruction ("Put the frog"), both age groups looked at the two visual referents of the linguistic input (two frogs) to the same extent (and more than at the other objects). This suggests that already children, like adults, identify multiple visual stimuli representing the meaning of a given linguistic input and integrate them in parallel into sentence processing.

Whether this holds true in the context of prediction—that is, whether children and adults anticipatorily fixate multiple visual prediction options in parallel when predicting input—was shown, for adults at least, by Ankener et al. (2018, Experiment 4) with a complex visual world paradigm. In an eye-tracking study, they auditorily presented adults with sentences containing semantically constraining verbs (e.g., "The man *spills* now the water) as participants looked at complex visual scenes of four objects. The novelty of this study was that the scenes varied in predictability; across four conditions, either 0, 1, 3, or 4 objects were consistent with the verb's semantic constraints, whereas the other objects presented on-screen were not (4, 3, 1, and 0, respectively) and therefore were considered as distractors. Notably, when inspecting adults' anticipatory fixations of the objects in the visual scenes, the authors found first evidence for a multiple predictions pattern. After hearing the constraining verb (spill) and before hearing the target word (water), adults anticipatorily fixated the target object (water) most frequently in the highly predictive 1-consistent condition, where only one object on-screen was consistent with the verb constraints, whereas the other three objects were distractors (cf. Andreu et al., 2013; Borovsky et al., 2012, Nation et al., 2003). In the less predictive 3- and 4-consistent conditions, in contrast, adults fixated the target object less often upon hearing the verb because they anticipatorily also fixated the two (lemonade, soup) or three (bowl, lemonade, soup) competitors that were also consistent with the verb constraints. These results suggest that adults can integrate more than one visual prediction option when anticipating upcoming linguistic input and thus follow a multiple predictions approach rather than a one-only approach to prediction in the complex visual world.

This conclusion was supported by the main finding of Ankener et al. (2018, Experiment 4), which was on the cognitive effort adults required to process the sentences in the different visual conditions (measured with the *index of cognitive activity*). The authors showed that adults' cognitive effort to process the predictable sentences was lower when they were presented with very predictive visual contexts showing only one visual prediction option (1-consistent) than when they were presented with multiple visual prediction options (3- and 4-consistent). This indicates that predictive sentence processing in adults is actually related to the predictability provided by the visual context and strengthens the finding that adults may have integrated one versus multiple visual prediction options across the different visual conditions.

In the current study, we sought to extend these findings in two ways. First, we aimed at examining whether it also holds true for children that they adapt their prediction behavior to the predictability of the visual contexts in such a way that they integrate either one or multiple visual prediction options. Second, we aimed at determining the role of individual differences in vocabulary size in the prediction of sentence input in such complex visual scenarios (see more details below). Given that the index of cognitive activity, in contrast to object fixations, is not an established measure in developmental research, we focused on the analysis of object fixations.

So far, it is unclear whether prediction also follows a multiple predictions pattern during early childhood. First evidence for this view was provided by Mani et al. (2016). They presented 2-year-olds with semantically constraining sentences (e.g., "The boy *reads* something") and visual scenes of two objects, both matching the verb constraints. Notably, the target object (e.g., book) was rated to be more strongly related to the verb than the competitor object (e.g., letter). Although children generally fixated the target more often than the competitor after hearing the verb, this preference decreased the lower the rated strength of the association between the target and the verb. Given that the authors analyzed proportional fixations to the target relative to the competitor, it can be concluded that, in these cases, children also fixated the competitor. This is a first sign that children can consider two visual prediction options in parallel when anticipating language. Moreover, Borovsky et al. (2012) presented 3- to 10-year-olds with visual scenes that displayed four objects of which only two objects (e.g., ship and treasure) were consistent with the semantically constraining agents (e.g., pirate) of auditorily presented sentences (e.g., "The *pirate* chases the treasure"). Upon hearing the agents, children anticipatorily fixated both consistent objects to a similar extent and thus were able to maintain two visual prediction options in parallel. Although this also indicates that children can maintain two visual prediction options, it remains unclear how they predict language in more complex visual contexts with a varying number of objects that are consistent with the linguistic constraints. In this regard, the current study examined whether the findings of Ankener et al. (2018, Experiment 4) on adults' object fixations in the complex visual world can be extended to young children. Thus, we investigated whether children are able to identify and maintain multiple suitable prediction options when predicting language in more complex visual contexts where the number of visual prediction options varies. In doing so, we examined children's anticipatory fixations of semantically consistent visual referents across such complex visual contexts.

Prediction and vocabulary size

Numerous studies highlight a relation between children's prediction skills and their verbal abilities (Borovsky & Creel, 2014; Borovsky et al., 2012; Lew-Williams & Fernald, 2007; Mani & Huettig, 2012, 2014, see Pickering & Gambi, 2018, for an overview). Thus, children's skill at predicting upcoming linguistic input is positively associated with their vocabulary size. That is, children with larger vocabulary size show improved prediction in the form of faster and more anticipatory fixations to potential target objects in visual contexts than other children (Borovsky & Creel, 2014; Borovsky et al., 2012; Fernald et al., 2006, 2008; Lew-Williams & Fernald, 2007; Mani & Huettig, 2012, 2014). There are a number of explanations for such findings, for instance, that children who generate predictions could use the mismatch between their predictions and the input to update their vocabulary size (Chang et al., 2006) and that children who quickly generate predictions have a processing advantage, sparing up resources required for vocabulary acquisition (Gambi, 2021). Equally, increased language experience in general could foster stronger prediction skills due to larger knowledge of the associations between linguistic units (Bar, 2009), which might shape the positive correlation of children's vocabulary size and their speed and extent of anticipatory fixations (Mani & Huettig, 2012). Against this background, we also examined the extent to which children's prediction in the complex visual world is related to individual differences in vocabulary size.

The current study

The purpose of this study was threefold. First, we aimed at revealing similar patterns of anticipatory object fixations in adults as shown by Ankener et al. (2018, Experiment 4), indicating that adults follow a multiple predictions approach in the complex visual world. Second, we examined whether young children show a similar multiple predictions pattern as adults in complex visual contexts. Because young children anticipatorily fixate visual referents consistent with the linguistic constraints, given a single consistent visual referent (Andreu et al., 2013; Mani & Huettig, 2012, 2014; Nation et al., 2003; Reuter et al., 2021) or even two consistent visual referents (Borovsky et al., 2012, Mani et al., 2016), we hypothesized that young children will anticipatorily fixate multiple visual referents (i.e., more than one) that are semantically consistent with the linguistic input presented thus far. In other words, we predicted that even young children maintain multiple prediction options in parallel when processing language. Alternatively, multiple suitable referents may challenge children's capabilities to come up with predictions and to maintain prediction options. As a result, children might pick out only one visual referent as a potential candidate or none at all.

To examine this, we presented children (5–6 years)[2] and adults with an eye-tracking task that was a child-friendly adaptation of the paradigm of Ankener et al. (2018, Experiment 4). Participants listened to German sentences containing semantically constraining verbs (e.g., "The father *eats* now the waffle") while looking at visual scenes of four objects. Across four conditions, 0, 1, 3, or 4 objects were consistent with the verb constraints. In contrast to Ankener et al. (2018), we focused only on participants' object fixations.

We expected the following results pattern. In each of the four visual conditions, children and adults should fixate all four objects upon hearing the sentences' subject (e.g., The father). This is because the subject was unconstraining and did not allow for a discrimination of the visual scene. Next, upon hearing the constraining verb (e.g., eat), both age groups were expected to fixate the target object (e.g., waffle) more often than all other objects in the highly predictive 1-consistent condition given that here only the target object was consistent with the verb constraints. This would reflect the classic prediction effect often reported for children (e.g., Mani & Huettig, 2014) and adults (e.g., Altmann & Kamide, 1999). In the less predictive 3- and 4-consistent conditions, where either three (pizza, sausage, waffle) or four (pretzel, pizza, sausage, waffle) visual objects were consistent with the verb constraints, we expected both age groups to show fewer anticipatory target fixations than in the 1-consistent condition. This is because here listeners should also fixate some of the competitors, suggesting that children and adults can integrate multiple visual prediction options into anticipatory sentence processing.

To further reveal whether children's and adults' prediction behavior is related to the exact number of visual prediction options, we compared the 3- and 4-consistent conditions. Upon hearing the verb, both age groups may show more target fixations in the 3-consistent condition than in the 4-consistent condition. This is because anticipatory fixations in the 3-consistent condition may fall on all three prediction options (and rarely on the single distractor), whereas those in the 4-consistent condition may be distributed among all four prediction options. A results pattern like this would suggest that children and adults not only follow a multiple predictions pattern but even adapt their prediction behavior to the exact number of visual prediction options presented.

Finally, when hearing the target word (e.g., waffle), we expected children and adults to fixate the target object (e.g., waffle) as the single visual referent of the target word more often than all other objects in the 1-, 3-, and 4-consistent conditions. In the 0-consistent control condition, where none of the visual objects was consistent with the verb constraints or the target word, both age groups were expected to similarly fixate all four semantically inconsistent distractors across the whole sentence.

[2] As in similar studies (e.g., Andreu et al., 2013; Borovsky & Creel, 2014; Borovsky et al., 2012; Reuter et al., 2021), the child sample was aged 5 and 6 years to ensure that the children knew the meaning of all verbs and recognized all visual objects of the eye-tracking task while being illiterate, thereby differing in language experience from the adult sample.

Third, we examined the extent to which children's language abilities, indexed by their receptive vocabulary size, are associated with their prediction skills. In keeping with previous studies reporting a positive relation between language abilities and prediction skills (Borovsky & Creel, 2014; Borovsky et al., 2012; Fernald et al., 2006, 2008; Lew-Williams & Fernald, 2007; Mani & Huettig 2012, 2014; Nation et al., 2003), we expected that children's receptive vocabulary size will be positively associated with their prediction of upcoming input. Hence, larger receptive vocabulary size of children should be accompanied by more anticipatory fixations to single or multiple visual prediction options upon hearing the constraining verb of a sentence.

Method

Participants

The final sample consisted of 26 children (5–6 years, M_{age} = 5.74 years, SD = 0.51) and 37 adults (19–40 years, M_{age} = 24.19 years, SD = 4.53). Data from an additional 4 children were excluded due to problems with eye-tracker calibration. Parents and adult participants filled in a form about their children's age, mother tongue, and vision or hearing impairment. All participants were German native speakers with reported normal or corrected-to-normal vision and hearing. For children, we used the form to ensure that they were not literate and further asked whether they had some problems with language comprehension or general development. Adults were asked about their highest academic degree, about their current employment, and whether they had some challenges with language comprehension that could affect their outcome in the study. Table 1 provides an overview of the demographic data. Participants were recruited via newspapers and flyers. All parents and adult participants gave informed consent and received compensation of 10 Euro. The study was approved by a local ethics committee of Saarland University.

Cognitive tests

We applied two standardized tests to control for the cognitive abilities of our sample (see Appendix A for a description of the tests). In the Semantic Verbal Fluency Task, a measure of cognitive functioning (Bialystok & Poarch, 2017; Friesen et al., 2014; Nielsen & Waldemar, 2006; Rosselli et al., 2009; Tröger et al., 2019; Troyer et al., 1997), children's performance (M = 10.24, SD = 4.72, range = 1–22) was comparable to other typically developing children of the same age (cf. Prigatano et al., 2008; Tallberg et al., 2011; van der Elst et al., 2011). Adults showed slightly higher verbal fluency (M = 26.12, SD = 4.92, range = 11–38) than in other studies (cf. Martins et al., 2007; Rosselli et al., 2002; Troyer et al., 1997; Zimmermann et al., 2014). In the Color Naming Task, a child-friendly test of processing speed (Karbach et al., 2011; Kray et al., 2006; Vergilova et al., 2022), children (M = 28.83, SD = 7.02, range = 12–40) performed in line with other samples of young children (Karbach et al., 2011; Kray et al., 2006) but worse than adults (M = 70.63, SD = 9.16, range = 46–84) . This suggests that the cognitive capacities of our child sample were comparable to other typically developing young children participating in empirical studies.

Experimental task

The eye-tracking task was programmed with EyeLink Experiment Builder (Version 2.2.61, SR Research, 2019a). Binocular data were collected at 500 Hz via EyeLink 1000+ in remote mode. As recommended for remote mode (SR Research, 2017), the eye-tracker was positioned below the experimental computer at a distance of about 50 cm from the participants. Prior to testing, we calibrated the gaze of each participant using a nine-point calibration procedure in which an attention-getter appeared in every position of a 3 × 3 grid of calibration points.

The task took about 20 minutes. Participants sat in front of a computer and were instructed to look at all four objects of a visual scene for 2000 ms (see Fig. 1). Then, a prerecorded sentence was presented auditorily (mean length = 5401 ms). The visual scene remained on-screen during the sentence

Table 1
Demographic information of the sample.

	Adults	Children
Age in years	*M* = 24.19, *SD* = 4.53, range = 19–40	*M* = 5.74, *SD* = 0.51, range = 4.8–6.7
Gender	15 men, 22 women	13 boys, 13 girls
Reading skills	–	*n* = 26 none
Writing skills	–	*n* = 26 none
Language challenges	None	*n* = 5 sigmatism
Developing challenges	–	None
Academic degree	*n* = 32 Abitur[a], *n* = 3 bachelor's degree, *n* = 2 master's degree	-
Employment	*n* = 33 student, *n* = 1 computer scientist, *n* = 1 data analyst, *n* = 1 geriatric caregiver	*n* = 26 kindergarten child

[a] Abitur is the highest school degree in Germany.

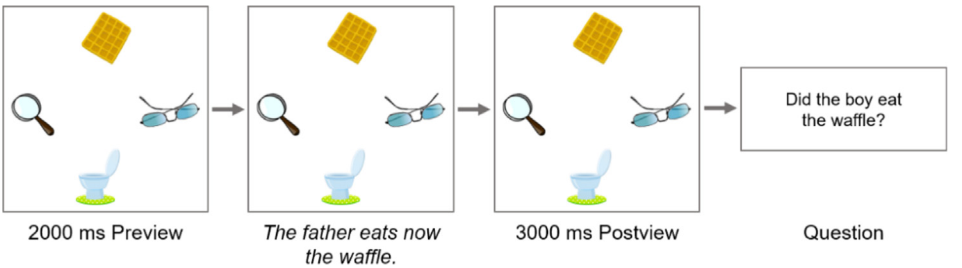

Fig. 1. Trial procedure. Example of a trial: Preview of the scene, followed by the auditorily presented sentence, the postview, and the question.

and for a postview of 3000 ms. Each trial was followed by a question on the sentence's subject, verb, or object (e.g., "Did the *boy* eat the waffle?") or the visual scene (e.g., "Was the waffle *on top* of the screen?"). Questions were correctly answered with either a "yes" or "no" response in 50% of cases. Adults read the questions on the screen and answered via button press on a button box. Children were presented auditorily with the prerecorded questions and were asked to respond verbally by saying "yes" or "no" to avoid potential issues related to the button press response at this age. The experimenter then input the button press response for them using the button box. The next trial started after a one-point calibration.

The experimental task consisted of 32 items, each made up of a six-word German sentence (e.g., "The father *eats* now the waffle") following the same syntactic structure (noun phrase – verb phrase – adverbial phrase – noun phrase). See Appendix B for all item sentences. The sentences were recorded by a female German native speaker slowly and in a voice appropriate for children using Audacity Version 2.3.3 (Audacity Team, 2019). The agents of the sentences were simple (e.g., father, mother), were uniformly distributed in terms of gender, and did not provide any clues for sentence continuation. The verbs were constraining (e.g., eat) and thus allowed only a limited number of arguments. "Now" was included as a spillover region to account for spillover from the verb to subsequent words during sentence presentation (cf. Ankener et al., 2018, Experiment 4). The objects were plausible verb arguments that were normed in a separate pretest with 40 children of the same age (Sommerfeld et al., 2022). Here, children were presented with colored pictures of the sentences' objects (e.g., waffle) next to three other objects. Children were first asked to name the object pictures to verify that they recognized them as intended (percentage of correct names per target object: *M* = 88.77, *SD* = 14.97, range = 50–100). Then, they were asked to identify those objects being consistent with the semantic constraints of a given verb (e.g., being *edible*). Plausibility ratings were at least 78% for each target object (M = 96.69, SD = 4.98, range = 78–100).

There were also 32 visual scenes (see supplementary material), each belonging to one sentence and each containing four colored object pictures arranged around the center (see Fig. 2). For each scene, the number of objects being consistent with the semantic verb constraints was manipulated such that either 0, 1, 3, or 4 objects were plausible verb arguments (see Fig. 2). For instance, a scene in the 3-consistent condition was made up of one target object (a picture of the sentence's object), two competitors (consistent with the verb constraints), and one distractor (inconsistent with the verb constraints). Plausibility ratings in the separate pretest were at least 70% for all target objects and competitors (M = 95.08, SD = 5.96, range = 70–100). Scenes were counterbalanced across the sentences in such a way that, for instance, a 0-consistent scene of one sentence served as a 4-consistent scene for a yoked sentence (see Fig. 2). This allowed us to control visual preferences for particular objects in the scene over and above the consistency of the objects given the sentence. Given that grammatical gender allows for article-based prediction in German (Bobb & Mani, 2013; Haeuser et al., 2020), only objects of the same grammatical gender were used within sentence pairs. Finally, the position of targets, competitors, and distractors was nearly completely rotated across items (completely balanced rotation was hindered because some items were removed after the pretest).

We used a latin square design to ensure that each sentence was presented in each of the four visual conditions (0-, 1-, 3-, and 4-consistent), whereas no participant experienced a sentence in more than one condition. This resulted in four different lists of 32 items each with 8 items per condition. Each participant was randomly assigned to one of the four lists. To mask the study design, we also presented participants with 8 filler trials introducing variation to the visual stimuli given that two objects were consistent with the verb constraints for each filler. The order of presentation of trials was randomized. Finally, we also included 3 practice trials at the beginning of the task. The filler and practice trials were the same for all participants.

Vocabulary assessment

We applied the Peabody Picture Vocabulary Test (Lenhard et al., 2015), which measures receptive vocabulary size (Borovsky & Creel, 2014; Borovsky et al., 2012; Lenhard et al., 2015; Vergilova et al., 2022). Participants were presented with a sheet of four colored pictures of nouns, verbs, or adjectives. Then, the experimenter vocalized the label of one of these nouns, verbs, or adjectives and asked participants to point at the picture representing the named item. There was no time limit, and participants could ask for the label to be repeated. They received no feedback (aside from 4 practice trials) and were asked to guess the answer if necessary. After the experimenter noted the answer, the next trial began. The test consisted of 19 trial sets, each consisting of 12 trials. The trial sets were arranged in order of ascending difficulty, and the participants' age determined which set was used to begin the test. If participants finished the last set or made eight mistakes or more within a set, the test was completed. The test lasted about 15 min. As in comparable studies (e.g., Borovsky & Creel, 2014; Vergilova et al., 2022), we extracted the raw Peabody Picture Vocabulary Test score, which was calculated as suggested in the manual: the number of the last trial minus the total number of mistakes. Higher scores indicate higher receptive vocabulary size.

Procedure

Each participant was tested alone in a 1-h session at Saarland University. After the consent form and questionnaire were filled in, participants completed the Semantic Verbal Fluency Task and Color Naming Task. To work on the experimental task, they moved to another desk with the eye-tracking setup. Here, they sat on a height-adjustable chair in front of a computer. The experimenter stuck a reference sticker on their forehead, as is required for remote eye-tracking. Successful camera adjustment and calibration were followed by written instructions for adults and verbal instructions for children. After participants completed the practice trials and were confirmed to understand the task, a divider was placed between participants and the experimenter to avoid distraction. Then, the experimental

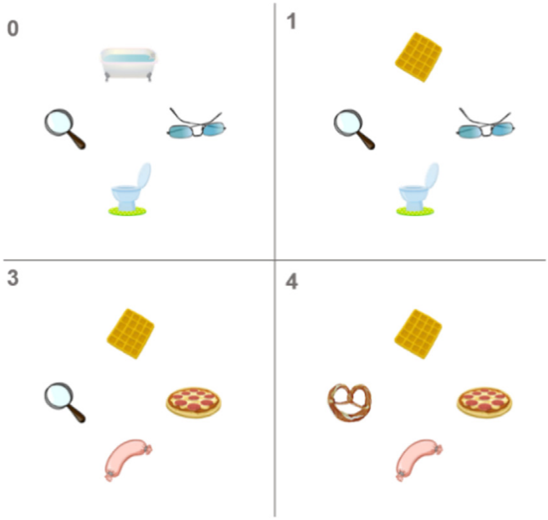

Fig. 2. Example of a visual scene in all four conditions. From left to right and from top to bottom, there are 0, 1, 3, or 4 consistent objects given the constraining sentence "The father *eats* now the waffle" (or 4, 3, 1, or 0 consistent objects given the counterbalanced sentence "The mother *cleans* now the magnifier").

task began. After completing the task, participants worked on the Peabody Picture Vocabulary Test and were finally compensated.

Data analysis

Fixations were extracted with EyeLink Data Viewer (SR Research, 2019b). Because all objects were the same size (650 × 650 mm) and appeared in the same four positions on the screen (see Fig. 2), we set up the same areas of interest for all visual scenes. For our analysis, we used only fixations that fell within these areas and excluded fixations shorter than 60 ms. Because the eye-tracker provided an estimate of where participants looked at each time stamp during the trial, with one data point approximately every 2 ms, we aggregated the data into 20-ms bins so that each 20-ms bin was coded for where participants were fixating. Because the onset of the different words varied between the items, we aggregated the 20-ms bins in three time windows separately for each participant and item. The baseline included all fixations within 2000 ms before the onset of the verb. The verb window included all fixations from 200 ms after verb onset until noun onset. The noun window included all fixations from 200 ms after noun onset until 2000 ms after noun onset. The delayed onset of the verb and noun window ensured that we included only fixations that can reliably be attributed to the auditory input in our analyses (Mani & Huettig, 2012, 2014; Mani et al., 2016).

The dependent variable of the experimental task was the proportion of fixations to the target relative to all other objects presented on-screen. This allowed us to show how fixations to the target and, due to the proportional character of this measure, to the other objects evolved across the trials. The target was defined as the pictorial representation of the object of each sentence. Because there was no target presented on-screen in the 0-consistent condition, we specified one distractor of this condition each as a pseudo-target in a counterbalanced way (i.e., the position of the pseudo-target was rotated across items and participants). Thus, we could use the proportion of target fixations as the dependent variable in the 0-consistent condition as well.

To be able to reveal whether children and adults can adapt their fixation behavior to the exact number of visual prediction options, we ran a control analysis in addition. We created a target advantage score (cf. Borovsky & Creel, 2014), which is defined as the difference between the proportion of fixations to the target minus each of the other objects in the visual scene. A score of 0 reflects an equal proportion of fixations to the target and the respective object. A positive score reflects a higher proportion of fixations to the target than the respective object. Hence, by comparing the target advantage scores of two objects, we can determine whether they were fixated to a similar or different extent. We computed three target advantage scores for each condition. In the 0- and 1-consistent conditions, the scores were defined as the difference between fixations to the (pseudo-)target and the first, second, or third distractor. In the 3-consistent condition, the scores were computed as the difference between fixations to the target and the first or second competitor or the single distractor. In the 4-consistent condition, the scores were defined as the difference between fixations to the target and the first, second, or third competitor.

Statistical analyses

All data were analyzed using RStudio Version 1.2.1335 (RStudio Team, 2018). A significance criterion of $p < .05$ was applied, and we report 95% confidence intervals of the mean. For all statistical analyses we applied linear mixed effects models because of their advantage to consider fixed and random effects in the experimental data at the same time (e.g., Baayen, 2008; Cunnings, 2012; Winter, 2013, 2020). These models do not only average data across participants or items when estimating the effects of the visual condition (0-, 1-, 3-, or 4-consistent targets) and the time window (baseline, verb, or noun) on the fixation behavior (as would be the case with more traditional analyses such as analysis of variance [ANOVA] designs). Instead, linear mixed effects models allow for the analysis of random effects and thus consider the variability in the data deriving from different participants and items. The "random intercepts" in the models take into account that multiple observations of a single participant across different items are not independent from each other (e.g., the mean proportion of target fixations could be 70% for one participant and 50% for another participant). Similar logic can be applied to different items. Moreover, the "random slopes" in the models account for the fact that different participants and items typically vary in their sensitivity to the visual manipulation (e.g., the proportion of target fixations could vary greatly depending on the visual condition for one participant and could vary only slightly for another participant). Hence, by including random intercepts and random slopes for participants and items in our statistical models, we were able to estimate any effect of the visual condition on the fixation behavior over and above the variability of the data across participants and items.

Results

We first provide an overview by describing the observed fixation pattern of both age groups across the four visual conditions and the three time windows (see Table 2 and Fig. 3). Then, we report the results of the analyses for the proportion of target fixations and then for the target advantage score. We also conducted additional post hoc analyses to underline that participants fixated multiple visual prediction options prior to hearing the target noun. Finally, we present the analysis of the relation between prediction and receptive vocabulary.

Description of the fixation pattern

As can be seen in Table 2 and Fig. 3, there were few differences in the fixation pattern of children and adults across the four visual conditions and time windows. In the baseline of all visual conditions, both age groups fixated all four objects. In the verb window of the 1-consistent condition, fixations to the single target object increased, whereas those to the three distractors decreased. In the verb window of the 3-consistent condition, fixations to the target object and the two competitors slightly increased, whereas those to the single distractor decreased. In the verb window of the 4-consistent condition, fixations fell on the target object and the three competitors. In the noun window of all conditions with a target object (1-, 3-, and 4-consistent), fixations to the target object increased, whereas those to all other objects decreased. In the 0-consistent condition, both groups‘ fixations fell at all four

Table 2
Averaged proportions of object fixations per condition, time window, and age group.

Condition	Group	Window	Object			
0			Pseudo-target	Distractor 1	Distractor 2	Distractor 3
	Adults	Baseline	.25 (.26)	.26 (.24)	.25 (.24)	.24 (.24)
		Verb	.27 (.22)	.24 (.22)	.26 (.22)	.23 (.21)
		Noun	.24 (.25)	.28 (.27)	.23 (.26)	.26 (.28)
	Children	Baseline	.25 (.26)	.23 (.24)	.27 (.26)	.25 (.26)
		Verb	.24 (.21)	.23 (.20)	.28 (.23)	.25 (.23)
		Noun	.25 (.27)	.21 (.23)	.29 (.27)	.25 (.25)
1			Target	Distractor 1	Distractor 2	Distractor 3
	Adults	Baseline	.25 (.26)	.26 (.25)	.25 (.25)	.25 (.26)
		Verb	.43 (.30)	.18 (.19)	.20 (.21)	.19 (.19)
		Noun	.58 (.36)	.15 (.23)	.15 (.24)	.12 (.20)
	Children	Baseline	.25 (.27)	.25 (.27)	.27 (.27)	.24 (.26)
		Verb	.44 (.29)	.19 (.19)	.17 (.19)	.20 (.20)
		Noun	.68 (.34)	.10 (.21)	.12 (.24)	.09 (.20)
3			Target	Competitor 1	Competitor 2	Distractor
	Adults	Baseline	.25 (.25)	.24 (.25)	.25 (.25)	.26 (.25)
		Verb	.27 (.23)	.25 (.21)	.31 (.25)	.17 (.20)
		Noun	.61 (.31)	.15 (.20)	.14 (.20)	.10 (.17)
	Children	Baseline	.23 (.25)	.25 (.27)	.24 (.24)	.28 (.28)
		Verb	.28 (.24)	.28 (.26)	.29 (.24)	.15 (.19)
		Noun	.60 (.30)	.15 (.21)	.18 (.23)	.08 (.13)
4			Competitor	Competitor 1	Competitor 2	Competitor 3
	Adults	Baseline	.24 (.23)	.24 (.24)	.27 (.26)	.25 (.25)
		Verb	.25 (.24)	.23 (.23)	.27 (.24)	.24 (.23)
		Noun	.57 (.31)	.14 (.18)	.14 (.19)	.15 (.21)
	Children	Baseline	.21 (.25)	.27 (.28)	.28 (.27)	.23 (.24)
		Verb	.24 (.24)	.27 (.25)	.24 (.24)	.25 (.24)
		Noun	.58 (.29)	.13 (.19)	.15 (.19)	.14 (.19)

Averaged across all participants and items, this table presents the proportions of adults' and children's fixations falling on each object of the visual scenes, that is, the (pseudo-)target, the distractors, and the competitors. Standard deviations are presented in parentheses.

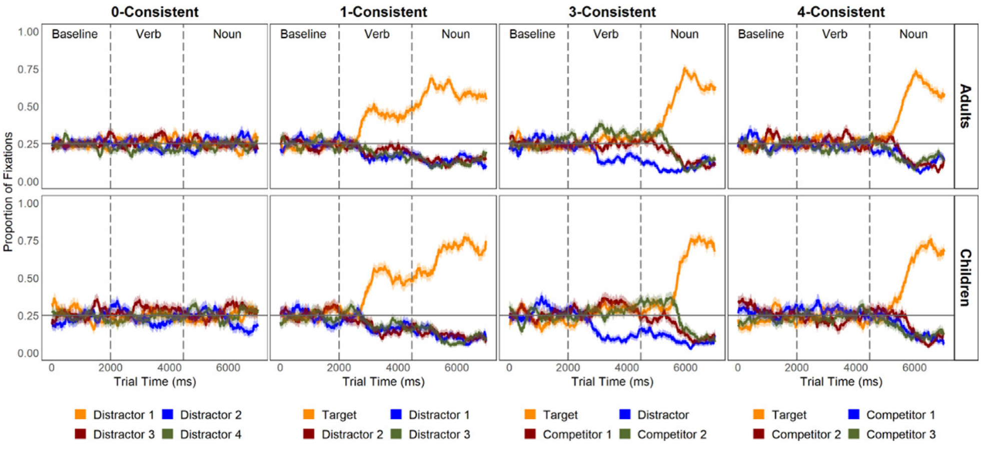

Fig. 3. Proportions of object fixations across the averaged trials. Adults' and children's proportions of fixations to the targets, distractors, and competitors in all conditions are shown. The baseline is shown in the first 2000 ms, the verb window between the two dashed lines, and the noun window following the second dashed line. Error bars indicate the standard error of the mean.

distractors across all time windows. The statistical analyses reported below confirm this descriptive pattern.

Analysis of the proportion of target fixations

We conducted a linear mixed effects model (lme4 library) on the proportion of target fixations with the factors condition (0-, 1-, 3-, or 4-consistent), time window (baseline, verb window, or noun window), and age group (adults or children). The factors were effect coded in a planned structure. For the factor condition, we defined three contrasts of most theoretical interest. First, we compared the 0-consistent condition with all other conditions. Second, we compared the 1-consistent condition with the 3- and 4-consistent conditions. Finally, we compared the 3-consistent condition with the 4-consistent condition. Time window contrasts were coded for two comparisons. First, we contrasted the baseline versus the verb window. Then, we contrasted the verb window versus the noun window. The age group contrast compared children with adults. These contrasts were added to the model including their interaction terms.

To consider variability across participants and items, the model structure included random intercepts for participants and items. Moreover, we included by-participant random slopes for the factors window and condition as well as by-item random slopes for the factors window, condition, and age group (see "Statistical analyses" section in Method). The interactions of the random slopes were also added. In case of a nonconverging model, it was simplified using the least-variance approach (Barr et al., 2013). The *p*-values were estimated using the Satterthwaite degrees of freedom method (lmerTest library), confidence intervals (*CI*s) with the stats library. For improved readability, we report only effects being relevant for our research questions. Appendix C shows the model with all results.[3]

Baseline versus verb window

There was an overall increase in proportional target fixations from the baseline to the verb window (β = −.22, *SE* = .01, *t*(5525.88) = −23.21, *p* < .001, *CI*[−.24, −.20]), which varied across conditions. There was a significant difference in the increase of proportional target fixations from the baseline to the verb window between the 0-consistent condition and all other conditions (β = .30, *SE* = .02, *t*(5525.38) = 13.59, *p* < .001, *CI*[.26, .34]), independent of the age group (*p* = .185). In the 0-consistent condition, there were no changes in both groups' fixations to the pseudo-target from the baseline (25%) to the verb window (26%, see Table 2). There also was a significant difference in the increase of proportional target fixations from the baseline to the verb window between the 1-consistent condition and the 3- and 4-consistent conditions (β = −.12, *SE* = .02, *t*(5527.76) = −5,35, *p* < .001, *CI*[−.17, −.08]), independent of the age group (*p* = .271). In the 1-consistent condition, both groups' proportional target fixations increased from 25% in the baseline to 43% in the verb window (see Table 2). In the 3- and 4-consistent conditions, the respective proportional target fixations changed from 24% and 23% in the baseline to 27% and 25% in the verb window. This small change in proportional target fixations from the baseline to the verb window did not vary between the 3- and 4-consistent conditions (*p* = .442) irrespective of the age group (*p* = .731).

Verb window versus noun window

There was an overall increase in the proportional target fixations from the verb to the noun window (β = −.32, *SE* = .02, *t*(75.09) = −21.03, *p* < .001, *CI*[−.35, −.29]), which varied across conditions. The proportional target fixations evolved significantly different from the verb to the noun window in the 0-consistent condition relative to all other conditions (β = .44, *SE* = .02, *t*(5530.62) = 20.02, *p* < .001, *CI*[.40, .49]) irrespective of the age group (*p* = .863). In the 0-consistent condition, both groups' proportional pseudo-target fixations did not change from the verb (26%) to the noun window (24%, see Table 2). In all other conditions, proportional target fixations increased from the verb to the noun window (see Fig. 3, orange line). Here, the 3- and 4-consistent conditions did not differ from each other

[3] We also computed the same model using only those items whose comprehension questions were answered correctly (i.e., we excluded 2 items for adults and 72 items for children). This did not change the pattern of results. Therefore, we report the model including all cases.

(p = .667) irrespective of the age group (p = .331). Both groups' proportional target fixations increased about 34% from the verb to the noun window in the 3-consistent (27% to 61%) and 4-consistent (25% to 58%) conditions (see Table 2). In the 1-consistent condition contrasted with the 3- and 4-consistent conditions, the proportion of target fixations increased significantly less (19%) from the verb (43%) to the noun (62%) window (β = .07, SE = .02, t(5530.10) = 3.05, p < .001, CI[.03, .12]). This difference was modulated by age group (β = .11, SE = .05, t(5529.99) = 2.40, p = .017, CI[.02, .20]). Post hoc linear mixed effect models with the same maximal random slope structure as in the last model separately for each age group (see Appendix D) revealed the following. Adults' proportional target fixations increased significantly more in the 3-consistent (34%) and 4-consistent (32%) conditions compared with the 1-consistent condition (15%) from the verb to the noun window (β = .13, SE = .03, t(3124.93) = 4.46, p < .001, CI[.07, .18]). Children's increase in target fixations from the verb to the noun window did not differ between the 3-consistent (32%) and 4-consistent (34%) conditions compared with the 1-consistent (24%) condition (p = .680).

In sum, there were no differences in children's and adults' anticipatory fixations of the target object. In the 1-, 3-, and 4-consistent conditions, both age groups fixated the target object as soon as the verb was played. However, they showed fewer anticipatory target fixations in the 1-consistent condition than in the 3- and 4-consistent conditions. Given the proportional character of the dependent variable, this suggests that children and adults must have also fixated some other objects upon hearing the verb in the 3- and 4-consistent conditions (see the next section for an analysis of which objects other than the target they fixated). Note that these results cannot be due to averaging data across participants (or items) given that our statistical models considered random effects of participants (and items).

Analysis of the target advantage score

To examine whether distractors and/or competitors were fixated to a similar or different extent across the trials — that is, whether participants adapted their prediction behavior to the exact number of visual prediction options — we analyzed the target advantage score (see "Data analysis" section in Method). Fig. 4 shows the averaged target advantage scores of both age groups across all conditions and time windows (see Appendix E for the numeric data). A small target advantage score means that the target and the respective object were fixated to a similar extent. A positive score means that the target was fixated more often than the respective distractor or competitor. A different score between the objects means that they were fixated to a different extent.

We ran one linear mixed effects model on the target advantage score for each of the four conditions (0-, 1-, 3-, and 4-consistent). The effect coded factors object, time window, and age group were included. Object contrasts were coded for two comparisons. We contrasted the first and second objects of a visual scene versus the third one. Then, we contrasted the first versus the second object. This was because in the 3-consistent condition both the first and second objects were competitors, whereas the third one was a distractor. Time window and age group contrasts were the same as in the analysis reported above. The contrasts were added to the models with their interaction terms. We included (a) random intercepts for participants and items and (b) by-participant random slopes for the factors object and window as well as by-item random slopes for the factors object, window, and age group. The interactions of the random slopes were added. Nonconverging models were simplified using the least-variance approach (Barr et al., 2013). All models are shown in Appendix F.

In the 0-, 1-, and 4-consistent conditions, the target advantage score did not differ among the three objects (p-values > .05) irrespective of the time window (p-values > .05) or age group (p-values > .05). This indicates that all three distractors in the 0- and 1-consistent conditions and all three competitors in the 4-consistent condition were fixated to a similar extent across the trials, that is, in the baseline, verb, and noun windows (see Fig. 4).

In the 3-consistent condition, the target advantage score did not vary between the two competitors (p = .240) independent of the time window (p-values > .05) and age group (p = .647). This shows that both groups fixated the two competitors to a similar extent across the trials in the 3-consistent condition. However, the score was significantly higher for the competitors than for the distractor (β = .06, SE = .01, t(4048.99) = 4.61, p < .001, CI[.03, .08]). This was unaffected by the age group (p = .523) but was modulated by the time window contrast comparing the baseline and the verb window (β = −.16,

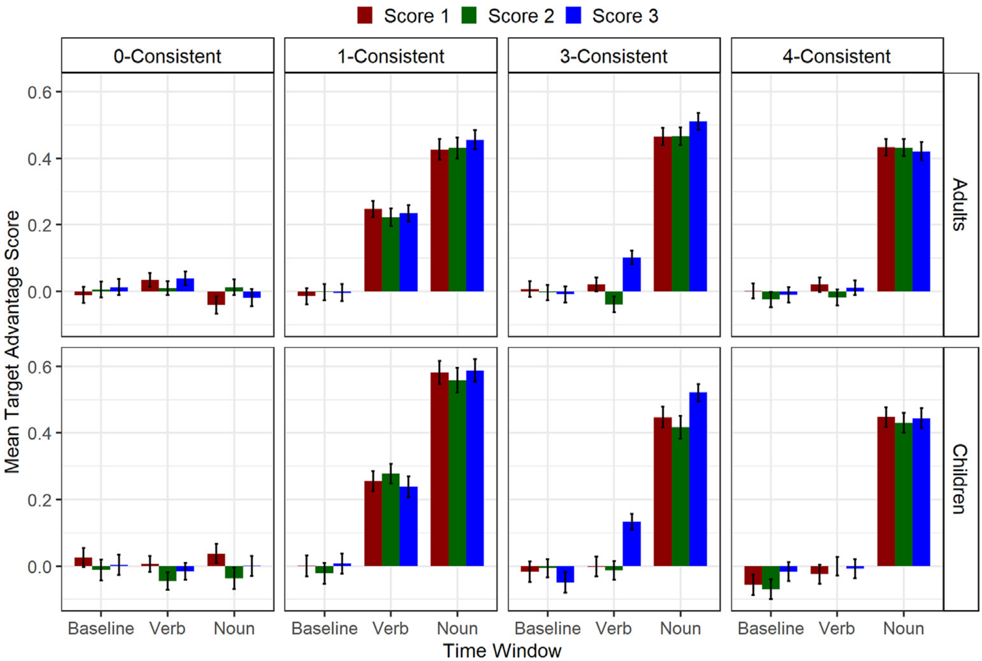

Fig. 4. Averaged target advantage scores per condition, time window, and age group. In the 0- and 1-consistent conditions, scores 1 to 3 reflect the difference of fixations on the (pseudo-)target minus each distractor. In the 3-consistent condition, scores 1 and 2 reflect the difference of fixations on the target minus each competitor and score 3 (blue bar) reflects the difference of fixations on the target minus the distractor. In the 4-consistent condition, scores 1 to 3 reflect the difference of fixations on the target minus each competitor. Error bars indicate the standard error of the mean.

SE = .03, $t(4048.99) = -4.63$, $p < .001$, $CI[-.23, -.09]$). Therefore, we ran post hoc linear mixed effects models with the same maximal random slope structure as in the last model separately for the baseline and the verb window (see Appendix G). In the baseline, the target advantage score did not differ among the competitors and the distractor (p = .284) irrespective of the age group (p = .558). In the verb window, the score was significantly lower for the competitors than for the distractor (β = .13, SE = .02, $t(1336.70) = 6.28$, $p < .001$, $CI[.09, .16]$) independent of the age group (p = .467). This results pattern suggests that both groups fixated the competitors and the distractor to a similar extent in the baseline. In the verb window, in contrast, the two competitors were fixated *more* often than the distractor and *as* often as the target. This was because the target advantage scores of the competitors were close to zero, whereas the score of the distractor was positive (see Fig. 4).

In sum, the analysis of the target average scores revealed no age differences in how children and adults fixated visual objects that are potential referents of the sentence context. Although we found no difference in the proportion of target fixations between the 3- and 4-consistent conditions, the analysis of the target advantage score provides evidence that both age groups adapted their prediction behavior to the number of visual prediction options. Upon hearing the verb, participants anticipatorily fixated the two competitors and the target object in the 3-consistent condition or all three competitors and the target object in the 4-consistent condition. These results were robust against variability across participants and items given that these effects were captured by the structure of our linear mixed effects models.

Post-hoc data analysis

To underline the results summarized above, we also manually determined the proportion of cases in which listeners fixated more than one/all visual prediction options in the verb window of the 3- and

4-consistent conditions. To do so, we first defined objects as "fixated" when at least 5% of all fixations of a given participant and trial fell on them. Then, we extracted the number of cases for which more than one/all visual prediction options were fixated and related this to the total number of cases. In the 3-consistent condition, adults anticipatorily fixated multiple (i.e., more than one) verb-consistent objects in 89% of all cases and children did so in 84% of all cases. In the 4-consistent condition, adults anticipatorily fixated multiple verb-consistent objects in 91% of all cases and children did so in 88% of all cases (see Table 3). Two chi-square tests showed that each of these percentages did not differ between the age groups (*p*-values > .05).

Second, we determined in how many cases listeners anticipatorily fixated all given visual prediction options. In the verb window of the 1-consistent condition, the target object was fixated by 86% of both the children and adults. In the 3-consistent condition, adults fixated all three visual prediction options (one target and two competitors) on hearing the verb in 50% of all cases and children did so in 47% of all cases. In the 4-consistent condition, adults anticipatorily fixated all four visual prediction options (one target and three competitors) in 36% of all cases and children did so in 31% of all cases (see Table 3). Three chi-square tests showed that each of these percentage values did not differ across the age groups (*p*-values > .05). See the supplementary material for a visualization of the anticipatory fixations of all objects across the visual conditions averaged across participants, averaged across items, and averaged separately for each combination of participant and item.

Influence of vocabulary size

Children's averaged raw score in the Peabody Picture Vocabulary Test (*M* = 116.62, *SD* = 24.86, range = 58–150) corresponded to a *t*-value of 52 (*CI*[47, 57]), as reported by the test manual for children aged 5.7 years (which was the mean age of our child sample). Thus, the receptive vocabulary size of our child sample was common for their age. Adults' raw scores in the Peabody Picture Vocabulary Test (*M* = 214.32, *SD* = 4.91, range = 201–224) corresponded to a *t*-value of 63 (*CI*[61, 65]), as reported by the manual for 17-year-olds (the German test version is only normed on 17 years). Thus, adults' receptive vocabulary size was above the norms of 17-year-olds. One child and one adult did not perform the test.

To analyze the interplay of prediction and receptive vocabulary size, the raw scores in the Peabody Picture Vocabulary Test were *z*-standardized separately for each age group and included in the following model. We ran a linear mixed effects model on the proportion of target fixations in the verb window of the 1- and 3-consistent conditions. This is because more fixations to the target than to the other objects in the verb window directly reflect prediction and because only the 1- and 3-consistent conditions allow for an anticipatory discrimination of the visual objects (in the 0- and 4-consistent conditions, either all or none of the objects are in line with prediction). We added the factors condition, age group, and receptive vocabulary size, together with their interactions, to the model. We also added random intercepts for participants and items as well as by-participant random slopes for the factor condition and by-item random slopes for the factors condition, age group, and receptive vocabulary size (including their interactions). In case of a nonconverging model, it was simplified with the least-variance approach (Barr et al., 2013). See Appendix H for the model results.

The model revealed a significant main effect of receptive vocabulary size (β = .03, *SE* = .01, *t*(61.11) = 2.76, *p* = .008, *CI*[.01, .05]), which was independent of the condition (*p* = .688) but was modulated by age group (β = −.05, *SE* = .02, *t*(63.06) = −2.42, *p* = .018, *CI*[−.09, −.01]). Post-hoc linear mixed effects models with the same maximal random slope structure as in the last model separately for each age group showed that children's anticipatory target fixations increased with increasing receptive vocabulary size (β = .05, *SE* = .01, *t*(26.72) = 4.14, *p* < .001, *CI*[.03, .08]). This was independent of the condition (*p* = .611) and was not the case for adults (*p* > .05). Fig. 5 visualizes this pattern of results.

Discussion

The main goals of this study were to examine age differences in predicting sentence continuations based on semantically constraining verbs in complex visual environments and whether the use of

Table 3
Exploratory data inspection.

Group	Condition	Observations	Multiple objects		All consistent objects	
			Number	Percentage	Number	Percentage
Adults	1	296	–	–	254	86
	3	296	264	89	148	50
	4	296	269	91	107	36
Children	1	199	–	–	171	86
	3	199	168	84	93	47
	4	198	175	88	62	31

Data address the verb window. "Observations" is the number of cases per condition across participants and items. The value 296 stems from 37 adults encountering 8 items per condition (37 × 8). For the 26 children who encountered 8 items per condition, there should be 208 cases per condition (26 × 8), but for some cases it was not possible to extract children's eye data. "Multiple objects" first shows the number of cases where more than one verb-consistent object was fixated. Next, this value is put in relation to the respective number of cases to get a percentage (e.g., 264 ÷ 296 is 89%). The cells of the 1-consistent condition are empty because here only the target object was consistent with the verb. "All consistent objects" shows the number and percentage of cases where *all* verb-consistent objects were fixated.

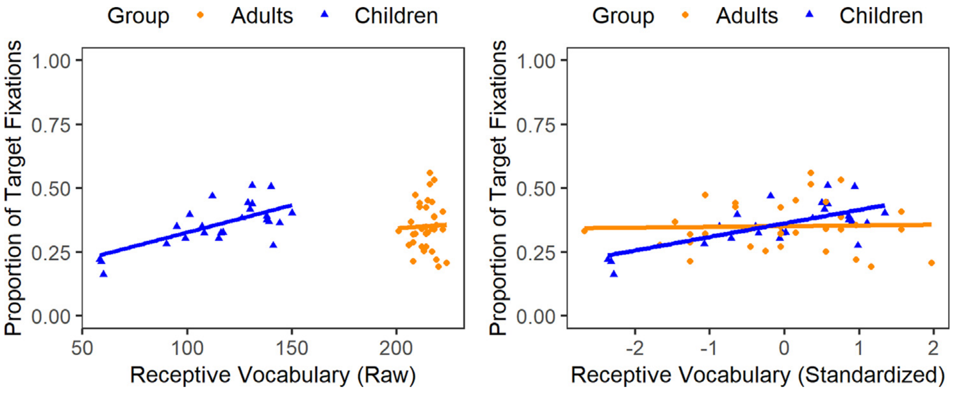

Fig. 5. Relation of anticipatory target fixations and receptive vocabulary size. Proportions of adults' and children's target fixations in the verb window of the 1- and 3-consistent conditions related to their raw (left panel) and *z*-standardized (right panel) Peabody Picture Vocabulary Test scores. Children's anticipatory fixations increased significantly with increasing receptive vocabulary.

prediction is associated with individual differences in vocabulary size. Therefore, we used eye-tracking in combination with a visual world paradigm and presented 5- and 6-year-old children and adults with sentences consisting of semantically constraining verbs while they were looking at complex visual scenes presenting four objects of which either 0, 1, 3, or 4 were consistent with the sentence context. Moreover, participants completed the Peabody Picture Vocabulary Test as a measure of receptive vocabulary size. In comparing children and adults in their fixation pattern across the experimental conditions, we found overall similarity but also some differences in the increase and decrease of fixations to the different objects across the sentences. Our results revealed four noteworthy findings that are discussed in more detail.

No age differences in predictive processing in complex visual scenes

The first finding is that children and adults predict sentence input even in complex visual contexts. This claim is supported by several results of our analysis. Considering first the predictive 1-, 3-, and 4-consistent conditions in which either one, three, or four visual objects were consistent with the sentence context. Here, children and adults first fixated all four objects during the nonpredictive beginning of the sentences (baseline). Upon hearing the constraining verb, they fixated the target object most frequently in the 1-consistent condition (whereas fixations to all three distractors decreased). In the 3- and 4-consistent conditions, both age groups fixated the target object less often than in the 1-consistent condition. This was because here they also fixated some of the other objects being consistent with the verb constraints (fixations to the single distractor decreased in the 3-consistent condition).

For the 1-consistent condition presenting one target next to multiple distractors, our finding is in line with the results of numerous eye-tracking studies in the visual world and replicates the classic prediction effect often reported for children and adults (cf. Altmann & Kamide, 1999; Andreu et al., 2013; Ankener et al., 2018; Borovsky & Creel, 2014; Borovsky et al., 2012; Mani & Huettig, 2014; Nation et al., 2003). Regarding the more novel 3- and 4-consistent conditions with multiple visual prediction options, our results are consistent with the fixation pattern of adults reported by Ankener et al. (2018, Experiment 4) but show that even young children also predict sentence input in complex visual environments.

Consider next the nonpredictive control condition in which none of the four visual objects was consistent with the sentence context. In the 0-consistent condition, children and adults fixated all four distractors across all time windows. This is also in line with the fixation-based findings of Ankener et al. (2018, Experiment 4) and shows that visual objects not consistent with linguistic constraints do not elicit children's and adults' anticipatory attention during sentence processing. In turn, this verifies that the fixation patterns in the predictive conditions (1-, 3-, and 4-consistent) were not due to chance but rather indicate that participants predicted the sentence input based on the constraining verbs in those other conditions.

Children and adults consider multiple prediction options

A second important finding is that even young children integrate multiple visual prediction options into anticipatory sentence processing. Upon hearing the verb, both age groups fixated one object when only one object was consistent with the verb constraints (1-consistent condition). They anticipatorily fixated multiple objects (i.e., more than one) when multiple objects were consistent with the verb constraints (3- and 4-consistent conditions). This can be concluded given that children and adults showed fewer anticipatory target fixations in the 3- and 4-consistent conditions than in the 1-consistent condition. Given that we analyzed the proportion of fixations to the target object relative to all other objects, this shows that both groups also fixated some of the other visual prediction options in the 3- and 4-consistent conditions. This suggests that children, like adults, integrate multiple visual prediction options into anticipatory sentence processing and argues against the possibility that they follow a one-only prediction fashion.

This finding presents a considerable advance on previous research that thus far has provided only few indications that children can integrate two visual cues into online sentence processing (Snedeker & Trueswell, 2004; Trueswell et al., 1999) and can integrate two visual prediction options into anticipatory sentence processing (Borovsky et al., 2012; Mani et al., 2016). The ability to track multiple prediction options may result from children's real-world experiences. Here, they experience language together with complex visual information that can even facilitate their linguistic processing (Knoeferle et al., 2005; van Rij, 2012; Weighall & Altmann, 2011; Zhang & Knoeferle, 2012) but also varies in its consistency with the linguistic input (Reuter et al., 2021). Such experiences may help children to learn to integrate multiple visual cues into language processing. Given that predicting upcoming input has positive associations with the speed of language processing, one option may be that children's maintaining of multiple options allows them to benefit from their predictions to a greater extent (Kuperberg & Jaeger, 2016; Mani & Huettig, 2014; van Petten & Luka, 2012). Given that the benefits of prediction are, moreover, thought to exceed its costs (Kuperberg & Jaeger, 2016), we conclude that already young children may be able to integrate much of the predictive information of complex visual contexts into language processing, thereby processing language efficiently in the complex visual world.

Consider next the question of whether children and adults can adapt their prediction behavior to the exact number of probable target objects in the visual scenes. Admittedly, we found no statistical difference in the proportion of target fixations between the 3- and 4-consistent conditions. However, our analysis on the target advantage scores revealed that upon hearing the verb in the 3-consistent condition, both age groups fixated the two competitors to a similar extent but more often than the single distractor. Here, the target advantage scores of the two competitors were close to zero, indicating that they were fixated to a similar extent as the target. In the verb window of the 4-consistent condition, the target advantage scores of all three competitors did not differ from each other and were close to zero. Hence, adults and children fixated all competitors and the target to a similar extent upon hearing the verb in this condition. These results indicate that both age groups can adapt their prediction behavior to the exact number of probable target objects in the visual scenes.

One may argue that the results in the 3- and 4-consistent conditions emerged from averaging data across participants and items. In principle, it could be the case that different participants fixated only one competitor, but each of them a different one on a particular item. Averaging across those participants would result in the same fixation pattern. The same argument can be made at the item level. However, we included random effects for participants and items into our models. Thus, all statistical results were revealed under the control of such participant- and item-specific variability, which is why we can rule out such an explanation for our data.

The post hoc data analysis further underlines our statistical results. Here, we manually determined the amount of cases in which participants anticipatorily fixated multiple or all visual prediction options. This inspection showed for the 3- and 4-consistent conditions that both age groups (a) anticipatorily fixated more than one visual prediction option in most cases and (b) adapted their prediction behavior to the exact number of visual prediction options for many cases (as also indicated by the results of our control analysis). This shows that both children and adults can follow a multiple predictions pattern in the complex visual world. Notably, the post hoc chi-square tests showed that the proportion of cases in which multiple or all visual prediction options were fixated anticipatorily did not differ across the age groups. This emphasizes again that young children's prediction behavior in the complex visual world is very similar to that of adults.

Children process target words differently than adults

Next, our study provides evidence of how children process dissolving sentence endings—that is, the target word at the end of a constraining sentence—when combined with complex visual contexts. Indeed, it has often been shown that children use the target word at the end of a sentence to guide their eye movements to the referent of this word in the visual scene (Andreu et al., 2013; Borovsky et al., 2012; Nation et al., 2003). This study revealed that they even do so when the given objects are potentially relevant for language processing at an earlier stage of the sentence. After the target was named, children's and adults' target fixations increased in the 3- and 4-consistent conditions, indicating that their competitor fixations now decreased as we analyzed the proportion of target fixations. This shows that even young children can inhibit previously integrated visual prediction options once a sentence is resolved and reveals how fast they can update their visual attention focus when processing language in the visual world.

However, the fixation patterns of children and adults differed slightly in the noun window. Adults' target fixations increased significantly less in the 1-consistent condition (15%) compared with the 3- and 4- consistent conditions (∼33%). Adults still paid more attention to the three distractors (∼14% each; see Table 2) than children (∼10% each) after the target was named in the 1-consistent condition. Children, in turn, fixated the target object almost exclusively here (∼69%). However, we note that children's increase in target fixations varied *descriptively* between the 1-consistent condition (24%) and the 3- and 4-consistent conditions (∼33%), although this difference was not statistically significant.

We explain this finding with recourse to the demands of language learning or inhibition of attention. Some accounts suggest that children use prediction to improve their language skills by comparing what they predicted with the actual input they received (Chang et al., 2006; Fazekas et al., 2020; Ramscar et al., 2013). Thus, they may have paid most attention to the fulfilled visual prediction option in the 1-consistent condition to memorize the connection between the semantically constraining verb and the semantical and perceptual properties of the target object. In contrast, adults, with their more advanced prediction and language comprehension skills, might not need to memorize the resolved combination of the linguistic and visual information. This would explain their reduced attention to the resolved target object in the 1-consistent condition.

Relations between prediction skills and vocabulary size

Finally, we also found evidence for age differences in the association between participants' prediction skills and vocabulary size. The Peabody Picture Vocabulary Test score was positively associated with children's—but not adults'—anticipatory target fixations. Children with larger receptive vocabularies showed increased anticipatory fixations to the target object in the 1- and 3-consistent conditions. For the 1-consistent condition, this is in line with results from a few other studies showing that the size of children's receptive vocabularies is associated with their anticipatory fixations toward single visual prediction options (Borovsky & Creel, 2014; Borovsky et al., 2012). For the 3-consistent condition, this suggests, for the first time, that children with larger receptive vocabularies also integrate complex visual contexts with more than one visual prediction option more strongly into predictive sentence processing than those with smaller vocabularies. This finding might be explained in terms of increased vocabulary knowledge itself being associated with increased knowledge of the associations between verbs and their arguments (Bar, 2009).

It should, however, be noted that the size of our child sample (n = 26) was relatively small to examine the influence of individual differences in vocabulary size on prediction. Our finding that children with larger vocabularies were more efficient at predicting how constraining sentences proceed than children with smaller vocabularies therefore should be treated with caution. However, we do note that, together with comparable studies (Borovsky & Creel, 2014; Borovsky et al., 2012), we provide further indication that children's language prediction may be positively related to their receptive vocabulary size.

Surprisingly, we did not find a modulation of prediction by receptive vocabulary size in our adult sample, although the studies by Borovsky and colleagues showed that adults' prediction in the visual world is also associated with their performance in the English version of the Peabody Picture Vocabulary Test (Borovsky & Creel, 2014; Borovsky et al., 2012). It is possible that this divergence is based on the restricted variance in the performance of our adult sample in this test given that the adult scores in the Peabody Picture Vocabulary Test were very homogeneous and in the upper range, hinting at a ceiling effect.

We do, however, note another limitation of the current study. We found that young children anticipatorily integrate more complex visual environments into sentence processing. Because the visual contexts varied in predictability, they were ecologically more valid than in other prediction research with children (cf. Andreu et al., 2013; Mani & Huettig, 2012, 2014; Nation et al., 2003). At the same time, it must be noted that four colored pictorial objects that are arranged in static visual scenes with systematically varying predictability are indeed closer to, but still far away from, the real world (cf. Reuter et al. 2021). Thus, this study provides a first clue that children are able to integrate complex visual contexts into predictive language processing, which remains to be investigated in more realistic and complex visual environments.

Conclusion

Our study revealed that young children (5–6 years), like adults, predict sentence input based on semantically constraining verbs even in complex visual contexts with more than one potential target object. In particular, we showed that children and adults can integrate multiple visual prediction options into anticipatory sentence processing, which in children was associated with receptive vocabulary size.

Data availability

The raw data and scripts for the analyses are available at https://doi.org/10.7910/DVN/UZJ0RZ.

Acknowledgment

This work was funded by the Deutsche Forschungsgemeinschaft (Project-ID: 232722074 – SFB 1102).

Author contributions

Linda Sommerfeld: conceptualization, data curation, formal analysis, investigation, methodology, project administration, software, supervision, validation, visualization, roles/writing–original draft, writing–review & editing; **Maria Staudte:** conceptualization, funding acquisition, methodology, project administration, resources, supervision, validation, writing–review & editing; **Nivedita Mani:** conceptualization, formal analysis, funding acquisition, methodology, validation, writing–review & editing; **Jutta Kray:** conceptualization, formal analysis, funding acquisition, methodology, project administration, resources, supervision, validation, writing–review & editing.

Appendix A

A.1. Semantic verbal fluency task

This task assesses semantic memory and executive functioning (Bialystok & Poarch, 2017; Friesen et al., 2014; Nielsen & Waldemar, 2006; Rosselli et al., 2009; Troyer et al., 1997) and thus can control the verbal and cognitive capacities of a sample. Participants were instructed to name all the animals coming to their mind within 60 s while avoiding repetitions. The verbal answers were recorded via Audacity (Version 2.3.2). Audio files were annotated in Praat (Version 6.0.37, Boersma & Weenink, 2019) by a German native speaker expert who transcribed correct answers (verified by https://de.wiktionary.org; last access December 9, 2021) and filtered out errors (e.g., forest) and disfluencies (e.g., hm). We automatically extracted the number of correct responses as the dependent variable using a custom Python (Version 3.7.3, Python Software Foundation, 2019) script. In total, 3 children and 1 adult did not perform this task.

A.2. Color naming task

This task measures perceptual processing speed and is a child-adapted version of the Digit Symbol Substitution Task (Karbach et al., 2011; Kray et al., 2006; Salthouse, 1992; Vergilova et al., 2022). Participants were presented with three sheets of paper on top of which they saw the same matching key consisting of four symbols in four colors (blue cross, green square, red triangle, and yellow circle). Below them were seven rows of these symbols, each containing all four objects in a different order but colorless. Thus, there were 28 symbols on each sheet for a total of 84 blank symbols. The matching key was presented on each sheet throughout the task. Participants were asked to name the color corresponding to each symbol referring to the key. They started with the first symbol on the first sheet and worked their way through line by line. Participants were asked to respond as quickly and accurately as possible within 60 s and were told that they might not reach the last symbol. They received no feedback (aside from 4 practice trials on an extra sheet with a different key). Responses were noted, and participants could correct errors. The number of correct responses minus the number of errors was calculated as the dependent variable (cf. Salthouse, 1992).

Appendix B

See Table B.1.

Table B.1
List of all item sentences of the experimental task.

	German sentence	English translation
1	Der Vater verschlingt gleich die Waffel.	The father eats now the waffle.
2	Die Mutter putzt gleich die Lupe.	The mother cleans now the magnifier.
3	Der Onkel schneidet gleich die Banane.	The uncle cuts now the banana.
4	Die Tante repariert gleich die Lampe.	The aunt repairs now the lamp.
5	Der Enkel probiert gleich die Erdbeere.	The grandson tastes now the strawberry.
6	Die Enkelin befüllt gleich die Gießkanne.	The granddaughter fills now the watering can.
7	Der Vater nascht gleich die Himbeere.	The father snacks now the raspberry.
8	Die Mutter entzündet gleich die Rakete.	The mother ignites now the rocket.
9	Der Großvater gießt gleich die Tomate.	The grandfather waters now the tomato.
10	Die Großmutter spielt gleich die Trompete.	The grandmother plays now the trumpet.
11	Der Bruder hört gleich die Gitarre.	The brother hears now the guitar.
12	Die Schwester kaut gleich die Pizza.	The sister chews now the pizza.
13	Der Bruder schließt gleich die Tasche.	The brother closes now the bag.
14	Die Schwester pflückt gleich die Kirsche.	The sister picks now the cherry.
15	Der Mann verschüttet gleich den Sprudel.	The man spills now the sparkling water.
16	Die Frau startet gleich den Computer.	The woman starts now the computer.
17	Der Opa backt gleich den Kuchen.	The granddad bakes now the cake.
18	Die Oma bremst gleich den Roller.	The grandma brakes now the scooter.
19	Der Mann erntet gleich den Salat.	The man harvests now the salad.
20	Die Frau näht gleich den Handschuh.	The woman sews now the glove.
21	Der Onkel sammelt gleich den Tannenzapfen.	The uncle collects now the fir cone.
22	Die Tante trinkt gleich den Kakao.	The aunt drinks now the cocoa.
23	Der Großvater zerbricht gleich den Pokal.	The grandfather breaks now the cup.
24	Die Großmutter strickt gleich den Schal.	The grandmother knits now the scarf.
25	Der Vater wirft gleich den Ball.	The father throws now the ball.
26	Die Mutter parkt gleich den Bus.	The mother parks now the bus.
27	Der Opa genießt gleich das Ei.	The granddad enjoys now the egg.
28	Die Oma wäscht gleich das Glas.	The grandma cleans now the glass.
29	Der Mann futtert gleich das Bonbon.	The man noshes now the candy.
30	Die Frau fährt gleich das Skateboard.	The woman rides now the skateboard.
31	Der Onkel baut gleich das Vogelhaus.	The uncle builds now the aviary.
32	Die Tante bügelt gleich das T-Shirt.	The aunt irons now the T-shirt.

In the experimental task, we presented the German sentences. The approximate English translations are provided for a better understanding of the semantic sentence constraints.

Appendix C

See Table C.1.

Table C.1
Results of the model on both age groups' proportions of target fixations.

Comparison	*β*	*SE*	*df*	*t*	*p*	*CI*
Intercept	.35	.01	61.62	30.75	.000	.33, .37
v1	−.14	.02	28.37	−8.33	.000	−.17, −.11
v2	.07	.02	45.84	4.87	.000	.04, .10
v3	.02	.01	5557.74	2.47	.014	.00, .04
t1	−.22	.01	5525.88	−23.21	.000	−.24, −.20
t2	−.32	.02	75.09	−21.03	.000	−.35, −.29
a	.00	.02	59.30	0.15	.882	−.03, .03
v1:t1	.30	.02	5525.38	13.59	.000	.26, .34
v1:t2	.44	.02	5530.62	20.02	.000	.40, .49
v2:t1	−.12	.02	5527.76	−5.35	.000	−.17, −.08
v2:t2	.07	.02	5530.10	3.05	.002	.03, .12
v3:t1	−.02	.03	5524.40	−0.77	.442	−.07, .03
v3:t2	−.01	.03	5526.68	−0.43	.667	−.06, .04
v1:a	.02	.02	5549.42	1.34	.181	−.01, .05
v2:a	−.05	.02	56.46	−1.97	.053	−.09, .00
v3:a	.00	.02	5551.95	−0.10	.922	−.04, .04
t1:a	.03	.02	5526.22	1.56	.119	−.01, .07
t2:a	.05	.03	75.08	1.55	.125	−.01, .11
v1:t1:a	−.06	.04	5525.57	−1.33	.185	−.15, .03
v1:t2:a	−.01	.04	5529.84	−0.17	.863	−.09, .08
v2:t1:a	.05	.05	5528.07	1.10	.271	−.04, .14
v2:t2:a	.11	.05	5529.99	2.40	.017	.02, .20
v3:t1:a	−.02	.05	5524.41	−0.34	.731	−.12, .09
v3:t2:a	−.05	.05	5526.61	−0.97	.331	−.16, .05

Results of the linear mixed effects model on the proportion of target fixations with the effect coded factors time window, condition, and age group.
Converged model: lmer(tf∼(v1 + v2 + v3)*(t1 + t2)*a+(1+(v2)*(t2)||participant) +(1+(v1 + v2)||item),data = all).
v1 to v3 are the condition contrasts, t1 and t2 are the time window contrasts, and a is the age group contrast.

Appendix D

See Table D.1.

Table D.1
Results of the models on the proportion of target fixations separately per age group.

Age group	Comparison	*β*	*SE*	*df*	*t*	*p*	*CI*
Adults	Intercept	.35	.01	51.67	27.04	.000	.32, .38
	v1	−.13	.02	45.88	−5.80	.000	−.17, −.08
	v2	.05	.01	27.08	3.61	.001	.02, .08
	v3	.02	.01	3124.31	1.97	.049	.00, .05
	t1	−.21	.02	35.43	−10.50	.000	−.25, −.17
	t2	−.30	.03	35.23	−10.95	.000	−.35, −.24
	v1:t1	.27	.04	34.23	6.65	.000	.19, .35
	v1:t2	.44	.05	33.42	8.83	.000	.34, .53
	v2:t1	−.10	.03	3122.58	−3.54	.000	−.16, −.04
	v2:t2	.13	.03	3124.93	4.46	.000	.07, .18
	v3:t1	−.03	.03	3121.06	−0.97	.332	−.10, .03
	v3:t2	−.04	.03	3121.75	−1.25	.213	−.11, .02

Table D.1 (*continued*)

Age group	Comparison	*β*	*SE*	*df*	*t*	*p*	*CI*
Children	Intercept	.35	.01	38.23	25.47	.000	.32, .38
	v1	−.15	.02	30.94	−6.88	.000	−.19, −.11
	v2	.10	.02	29.82	4.37	.000	.05, .14
	v3	.03	.01	2208.74	1.76	.078	.00, .06
	t1	−.24	.01	2159.79	−16.05	.000	−.27, −.21
	t2	−.35	.01	2160.06	−23.34	.000	−.37, −.32
	v1:t1	.33	.03	2161.27	9.65	.000	.26, .40
	v1:t2	.45	.03	2159.99	13.15	.000	.38, .52
	v2:t1	−.15	.04	2159.61	−4.19	.000	.32, .38
	v2:t2	.01	.04	2161.54	0.41	.680	−.17, −.08
	v3:t1	−.01	.04	2158.11	−0.26	.797	.02, .08
	v3:t2	.02	.04	2157.97	0.39	.694	.00, .05

Results of the linear mixed effects models on the proportion of target fixations with the effect coded factors condition and time window separately per age group.
Converged model adults: lmer(tf~(v1 + v2 + v3)*(t1 + t2)+(1+(v1)*(t1 + t2)||participant) +(1+(v1 + v2)||item),data = adults).
Converged model children: lmer(tf~(v1 + v2 + v3)*(t1 + t2)+(1+(v1 + v2)||participant) +(1+(v1 + v2)||item),data = children).
v1 to v3 are the condition contrasts, and t1 and t2 are the time window contrasts.

Appendix E

See Table E.1.

Table E.1
Averaged target advantage scores per condition, time window, and age group.

Condition	Age group	Window	Object		
0			Distractor 1	Distractor 2	Distractor 3
	Adults	Baseline	−.01 (.42)	.00 (.41)	.01 (.42)
		Verb	.04 (.36)	.01 (.36)	.04 (.36)
		Noun	−.04 (.42)	.01 (.39)	−.02 (.43)
	Children	Baseline	.03 (.39)	−.01 (.43)	.00 (.42)
		Verb	.01 (.33)	−.05 (.38)	−.02 (.35)
		Noun	.04 (.4)	−.04 (.45)	.00 (.42)
1			Distractor 1	Distractor 2	Distractor 3
	Adults	Baseline	−.01 (.42)	.00 (.41)	.00 (.44)
		Verb	.25 (.42)	.22 (.45)	.24 (.42)
		Noun	.43 (.52)	.43 (.53)	.46 (.49)
	Children	Baseline	.00 (.45)	−.02 (.44)	.01 (.42)
		Verb	.26 (.41)	.28 (.40)	.24 (.44)
		Noun	.58 (.49)	.56 (.51)	.59 (.47)
3			Competitor 1	Competitor 2	Distractor
	Adults	Baseline	.01 (.40)	.00 (.39)	−.01 (.42)
		Verb	.02 (.36)	−.04 (.41)	.10 (.35)
		Noun	.47 (.45)	.47 (.45)	.51 (.43)
	Children	Baseline	−.02 (.43)	−.01 (.39)	−.05 (.43)
		Verb	.00 (.42)	−.01 (.39)	.13 (.33)
		Noun	.45 (.44)	.42 (.48)	.52 (.36)
4			Competitor 1	Competitor 2	Competitor 3
	Adults	Baseline	.00 (.38)	−.02 (.39)	−.01 (.39)
		Verb	.02 (.37)	−.02 (.40)	.01 (.39)
		Noun	.43 (.42)	.43 (.44)	.42 (.46)
	Children	Baseline	−.06 (.43)	−.07 (.42)	−.02 (.40)
		Verb	−.03 (.40)	.00 (.38)	−.01 (.40)
		Noun	.45 (.41)	.43 (.42)	.44 (.42)

The "Object" column shows which object was referenced to the (pseudo-)target.
Standard deviations are presented in parentheses.

Appendix F

See Table F.1.

Table F.1
Results of the models on the target advantage score.

Condition	Comparison	*β*	*SE*	*df*	*t*	*p*	*CI*
0	Intercept	.00	.02	55.88	−0.24	.811	−.04, .03
	o1	.00	.01	4159.47	0.38	.703	−.02, .03
	o2	−.02	.02	60.22	−1.12	.267	−.06, .02
	t1	.01	.02	4163.16	0.25	.800	−.03, .04
	t2	.02	.02	4163.02	0.94	.347	−.02, .06
	a	.03	.03	65.48	0.87	.386	−.03, .08
	o1:t1	.00	.03	4159.47	0.12	.906	−.06, .07
	o1:t2	.01	.03	4159.47	0.41	.680	−.05, .08
	o2:t1	.02	.04	4164.25	0.44	.660	−.06, .10
	o2:t2	−.02	.04	4164.73	−0.49	.627	−.10, .06
	o1:a	.01	.02	4159.47	0.36	.719	−.04, .06
	o2:a	.07	.04	60.22	1.93	.059	.00, .15
	t1:a	−.01	.04	4163.01	−0.24	.807	−.09, .07
	t2:a	.06	.04	4163.74	1.60	.110	−.01, .14
	o1:t1:a	.02	.07	4159.47	0.29	.771	−.12, .16
	o1:t2:a	.03	.07	4159.47	0.40	.688	−.11, .16
	o2:t1:a	−.03	.08	4164.25	−0.37	.711	−.19, .13
	o2:t2:a	−.12	.08	4164.73	−1.47	.142	−.27, .04
1	Intercept	.24	.03	67.07	8.14	.000	.18, .30
	o1	.01	.01	3977.66	0.52	.603	−.02, .03
	o2	−.01	.01	3977.66	−0.35	.725	−.03, .02
	t1	−.50	.04	79.22	−11.20	.000	−.58, −.41
	t2	−.52	.05	81.84	−9.96	.000	−.62, −.42
	a	−.04	.05	81.21	−0.83	.408	−.13, .05
	o1:t1	.01	.03	3977.66	0.27	.785	−.06, .08
	o1:t2	−.03	.04	3977.66	−0.89	.372	−.10, .04
	o2:t1	.00	.04	3977.66	0.02	.986	−.08, .08
	o2:t2	.01	.04	3977.66	0.22	.830	−.07, .09
	o1:a	.01	.02	3977.66	0.32	.751	−.04, .06
	o2:a	.00	.03	3977.66	0.17	.863	−.05, .06
	t1:a	.08	.09	74.09	0.93	.357	−.09, .25
	t2:a	.16	.10	79.16	1.64	.106	−.03, .35
	o1:t1:a	−.04	.07	3977.66	−0.61	.544	−.18, .09
	o1:t2:a	.00	.07	3977.66	−0.04	.971	−.14, .13
	o2:t1:a	.06	.08	3977.66	0.73	.464	−.10, .22
	o2:t2:a	−.04	.08	3977.66	−0.55	.579	−.20, .11
3	(Intercept)	.18	.02	66.49	10.84	.000	.15, .22
	o1	.06	.01	4048.99	4.61	.000	.03, .08
	o2	−.02	.01	4048.99	−1.17	.240	−.04, .01
	t1	−.40	.03	76.95	−12.36	.000	−.47, −.34
	t2	−.62	.04	83.08	−13.79	.000	−.71, −.53
	a	.01	.03	74.06	0.19	.850	−.05, .06
	o1:t1	−.16	.03	4048.99	−4.63	.000	−.23, −.09
	o1:t2	−.02	.03	4048.99	−0.64	.526	−.09, .05
	o2:t1	.03	.04	4048.99	0.86	.392	−.04, .11
	o2:t2	.00	.04	4048.99	−0.09	.927	−.08, .07
	o1:a	−.02	.02	4048.99	−0.64	.523	−.06, .03
	o2:a	−.01	.03	4048.99	−0.46	.647	−.07, .04
	t1:a	.04	.07	76.95	0.63	.530	−.09, .17
	t2:a	−.01	.08	73.42	−0.09	.927	−.17, .15

Table F.1 (*continued*)

Condition	Comparison	*β*	*SE*	*df*	*t*	*p*	*CI*
	o1:t1:a	.08	.07	4048.99	1.20	.231	−.05, .22
	o1:t2:a	.06	.07	4048.99	0.81	.419	−.08, .19
	o2:t1:a	−.02	.08	4048.99	−0.21	.835	−.17, .14
	o2:t2:a	−.09	.08	4048.99	−1.09	.277	−.24, .07
4	Intercept	.13	.02	62.53	6.66	.000	.09, .17
	o1	.01	.01	3990.17	0.77	.442	−.01, .03
	o2	−.01	.01	3990.17	−0.84	.399	−.04, .02
	t1	−.32	.03	74.48	−1.46	.000	−.37, −.26
	t2	−.59	.04	76.56	−16.13	.000	−.66, −.52
	a	.01	.03	74.33	0.43	.670	−.05, .08
	o1:t1	.03	.03	3990.17	0.83	.404	−.04, .10
	o1:t2	.03	.03	3990.17	0.72	.470	−.04, .09
	o2:t1	−.02	.04	3990.17	−0.38	.703	−.09, .06
	o2:t2	−.01	.04	3990.17	−0.15	.878	−.08, .07
	o1:a	−.02	.02	3990.17	−0.79	.428	−.07, .03
	o2:a	−.02	.03	3990.17	−0.69	.494	−.07, .04
	t1:a	.03	.06	80.29	0.50	.621	−.08, .14
	t2:a	.05	.07	76.56	0.72	.475	−.09, .20
	o1:t1:a	−.05	.07	3990.17	−0.76	.446	−.19, .08
	o1:t2:a	.00	.07	3990.17	−0.07	.944	−.14, .13
	o2:t1:a	.02	.08	3990.17	0.19	.847	−.14, .17
	o2:t2:a	−.07	.08	3990.17	−0.90	.371	−.23, .09

Results of the linear mixed effects models on the target advantage score with the effect coded factors object, time window, and age group.
The converged models per condition are as follows:
Condition 0: lmer(tas~(o1 + o2)*(t1 + t2)*a+(1+(o2)||participant)+(1 + a||item),data = cond_0).
Condition 1: lmer(tas~(o1 + o2)*(t1 + t2)*a+(1+(t1 + t2)||participant)+(1+(t1 + t2)*a||item),data = cond_1).
Condition 3: lmer(tas~(o1 + o2)*(t1 + t2)*a+(1+(t1 + t2)||participant)+(1+(t2)*a||item),data = cond_3).
Condition 4: lmer(tas~(o1 + o2)*(t1 + t2)*a+(1+(t1 + t2)||participant)+(1+(t1)*a||item),data = cond_4).
o1 and o2 indicate the first and second object contrasts, t1 and t2 indicate the first and second time window contrasts, and a indicates the age group contrast.

Appendix G

See Table G.1.

Table G.1
Results of the models on the target advantage score in the 3-consistent condition in the baseline and verb window.

Window	Comparison	*β*	*SE*	*df*	*t*	*p*	*CI*
Baseline	Intercept	−.02	.02	52.44	−0.99	.325	−.06, .02
	o1	−.02	.02	1345.09	−1.07	.284	−.07, .02
	o2	.00	.03	1345.09	0.02	.985	−.05, .05
	a	.03	.04	52.81	0.66	.509	−.06, .11
	o1:a	.03	.04	1345.09	0.59	.558	−.06, .11
	o2:a	−.02	.05	1345.09	−0.42	.676	−.12, .08
Verb	Intercept	.08	.02	48.57	3.27	.002	.03, .12
	o1	.13	.02	1336.70	6.28	.000	.09, .16
	o2	−.04	.02	1336.70	−1.54	.124	−.08, .01
	a	−.02	.04	56.33	−0.57	.569	−.09, .05
	o1:a	−.03	.04	1336.70	−0.73	.467	−.11, .05
	o2:a	−.05	.05	1336.70	−1.05	.296	−.14, .04

Results of the two separate linear mixed effects models on the target advantage score in the 3-consistent condition in the baseline and verb windows with the effect coded factors object and age group.
The converged models per time window are as follows:
Baseline: lmer(tas~(o1 + o2)*a+(1|participant)+(1 + a||item),data = cond_3_base).
Verb window: lmer(tas~(o1 + o2)*a+(1|participant)+(1 + a||item),data = cond_3_verb).
o1 and o2 indicate the first and second object contrasts, and a indicates the age group contrast.

Appendix H

See Table H.1.

Table H.1
Results on the interplay of prediction and receptive vocabulary size.

Age group	Comparison	*β*	*SE*	*df*	*t*	*p*	*CI*
Overall	Intercept	.36	.02	40.08	21.16	.000	.32, .39
	v	.16	.02	37.46	6.87	.000	.11, .20
	a	−.01	.02	56.80	−0.42	.674	−.05, .03
	ppvtZ	.03	.01	61.11	2.76	.008	.01, .05
	v:a	−.01	.04	54.21	−0.18	.860	−.09, .07
	v:ppvtZ	.01	.02	58.74	0.40	.688	−.03, .05
	a:ppvtZ	−.05	.02	63.06	−2.42	.018	−.09, −.01
	v:a:ppvtZ	−.01	.04	61.04	−0.31	.757	−.09, .07
Adults	Intercept	.35	.02	39.24	17.62	.000	.31, .39
	v	.16	.03	29.69	5.09	.000	.10, .22
	ppvtZ	.00	.02	34.73	0.23	.823	−.03, .03
	v:ppvtZ	.00	.03	33.80	0.05	.960	−.05, .06
Children	Intercept	.36	.02	22.65	19.81	.000	.33, .40
	v	.16	.03	15.57	5.54	.000	.11, .22
	ppvtZ	.05	.01	26.72	4.14	.000	.03, .08
	v:ppvtZ	.01	.03	25.64	0.51	.611	−.04, .07

Results of the linear mixed effects models on the proportion of target fixations with the factors condition, age group, and receptive vocabulary size.
Converged model overall: lmer(tar ~ c*a*ppvtZ+(1 + c||participant)+(1 + c||item),data = all).
Converged model adults: lmer(tar ~ c*ppvtZ+(1 + c||participant)+(1 + c||item),data = adults).
Converged model children: lmer(tar ~ c*ppvtZ+(1 + c||participant)+(1 + c||item),data = children).
v is the condition, a is the age group, and ppvtZ is the receptive vocabulary size factor.

Appendix I. Supplementary data

Supplemental data, i.e., the visual scenes of the experimental task (Visual contexts.pptx) and the visualized fixation patterns per participant and trial (Fixations per participant & trial.7z), can be found online at https://doi.org/10.1016/j.jecp.2023.105690. All supplemental material, including the experimental stimuli as well as the data and scripts of the analyses, is provided online at https://doi.org/10.7910/DVN/UZJ0RZ.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*(4), 419–439. https://doi.org/10.1006/jmla.1997.2558.

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247–264. https://doi.org/10.1016/S0010-0277(99)00059-1.

Andreu, L., Sanz-Torrent, M., & Trueswell, J. C. (2013). Anticipatory sentence processing in children with specific language impairment: Evidence from eye movements during listening. *Applied Psycholinguistics, 34*(1), 5–44. https://doi.org/10.1017/S0142716411000592.

Ankener, C. S., Sekicki, M., & Staudte, M. (2018). The influence of visual uncertainty on word surprisal and processing effort. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.02387. Article 2387.

Audacity Team. (2019). *Audacity, die kostenlose, quelloffene, plattformübergreifende Software zur Aufnahme und Bearbeitung von Klängen (Version 2.3.2)* [computer software]. Audacity. https://www.audacityteam.org.

Baayen, H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.

Bar, M. (2009). The pro-active brain: Memory for predictions. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1521), 1235–1243. https://doi.org/10.1098/rstb.2008.0310.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001.

Bobb, S. C., & Mani, N. (2013). Categorizing with gender: Does implicit grammatical gender affect semantic processing in 24-month-old toddlers? *Journal of Experimental Child Psychology, 115*(2), 297–308. https://doi.org/10.1016/j.jecp.2013.02.006.

Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer (Version 6.0.37)* [computer software]. https://www.praat.org.
Borovsky, A., & Creel, S. C. (2014). Children and adults integrate talker and verb information in online processing. *Developmental Psychology, 50*(5), 1600–1613. https://doi.org/10.1037/a0035591.
Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology, 112*(4), 417–436. https://doi.org/10.1016/j.jecp.2012.01.005.
Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*(2), 234–272. https://doi.org/10.1037/0033-295X.113.2.234.
Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research, 28*(3), 369–382. https://doi.org/10.1177%2F0267658312443651.
Engelhardt, P. E., Ferreira, F., & Patsenko, E. G. (2010). Pupillometry reveals processing load during spoken language comprehension. *Quarterly Journal of Experimental Psychology, 63*(4), 639–645. https://doi.org/10.1080/17470210903469864.
Fazekas, J., Jessop, A., Pine, J., & Rowland, C. (2020). Do children learn from their prediction mistakes? A registered report evaluating error-based theories of language acquisition. *Royal Society Open Science, 7*(11). https://doi.org/10.1098/rsos.180877 180877.
Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology, 42*(1), 98–116. https://doi.org/10.1037/0012-1649.42.1.98.
Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), Language acquisition and language disorders, Vol. 44: Developmental psycholinguistics: On-line methods in children's language processing (pp. 97–135). John Benjamins. https://doi.org/10.1075/lald.44.06fer.
Gambi, C. (2021). The role of prediction in second language vocabulary learning. In E. Kaan & T. Grüter (Eds.). *Prediction in second language processing and learning* (Vol. 12, pp. 187–206). John Benjamins. https://doi.org/10.1075/bpa.12.09gam.
Haeuser, K. I., Kray, J., & Borovsky, A. (2020, July–August). *Great expectations: Evidence for graded prediction of grammatical gender.* Paper presented at the 42nd annual meeting of the Cognitive Science Society, Seattle, WA.
Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137*(2), 151–171. https://doi.org/10.1016/j.actpsy.2010.11.003.
Karbach, J., Kray, J., & Hommel, B. (2011). Action–effect learning in early childhood: Does language matter? *Psychological Research, 75*(4), 334–340. https://doi.org/10.1007/s00426-010-0308-1.
Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition, 95*(1), 95–127. https://doi.org/10.1016/j.cognition.2004.03.002.
Kray, J., Eenshuistra, R., Kerstner, H., Weidema, M., & Hommel, B. (2006). Language and action control: The acquisition of action goals in early childhood. *Psychological Science, 17*(9), 737–741. https://doi.org/10.1111/j.1467-9280.2006.01774.x.
Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience, 31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299.
Lenhard, A., Lenhard, W., Suggate, S., Segerer, R. (2015). *Peabody Picture Vocabulary Test–Revision IV: Technical manual* [German adaptation]. Pearson Assessment.
Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 18*(3), 193–198. https://doi.org/10.1111/j.1467-9280.2007.01871.x.
Mahr, T., McMillan, B. T., Saffran, J. R., Weismer, S. E., & Edwards, J. (2015). Anticipatory coarticulation facilitates word recognition in toddlers. *Cognition, 142*, 345–350. https://doi.org/10.1016/j.cognition.2015.05.009.
Mani, N., Daum, M. M., & Huettig, F. (2016). "Proactive" in many ways: Developmental evidence for a dynamic pluralistic approach to prediction. *Quarterly Journal of Experimental Psychology, 69*(11), 2189–2201. https://doi.org/10.1080%2F17470218.2015.1111395.
Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—but only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 843–847 https://psycnet.apa.org/doi/10.1037/a0029284.
Mani, N., & Huettig, F. (2014). Word reading skill predicts anticipation of upcoming spoken language input: A study of children developing proficiency in reading. *Journal of Experimental Child Psychology, 126*, 264–279. https://doi.org/10.1016/j.jecp.2014.05.004.
Martins, I. P., Vieira, R., Loureiro, C., & Santos, M. E. (2007). Speech rate and fluency in children and adolescents. *Child Neuropsychology, 13*(4), 319–332. https://doi.org/10.1080/09297040600837370.
Nation, K., Marshall, C. M., & Altmann, G. T. (2003). Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. *Journal of Experimental Child Psychology, 86*(4), 314–329. https://doi.org/10.1016/j.jecp.2003.09.001.
Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin, 144*(10), 1002–1044 https://psycnet.apa.org/doi/10.1037/bul0000158.
Prigatano, G. P., Gray, J. A., & Lomay, V. T. (2008). Verbal (animal) fluency scores in age/grade appropriate minority children from low socioeconomic backgrounds. *Journal of the International Neuropsychological Society, 14*(1), 143–147. https://doi.org/10.1017/S1355617708080089.
Python Software Foundation. (2019). *Python Programming Language (Version 3.7.3)* [computer software]. https://www.python.org.
Ramscar, M., Dye, M., & McCauley, S. M. (2013). Error and expectation in language learning: The curious absence of mouses in adult speech. *Language, 89*(4), 760–793. https://doi.org/10.1353/lan.2013.0068.
Reuter, T., Dalawella, K., & Lew-Williams, C. (2021). Adults and children predict in complex and variable referential contexts. *Language, Cognition and Neuroscience, 36*(4), 474–490. https://doi.org/10.1080/23273798.2020.1839665.
Rosselli, M., Ardila, A., Salvatierra, J., Marquez, M., Luis, M., & Weekes, V. A. (2002). A cross-linguistic comparison of verbal fluency tests. *International Journal of Neuroscience, 112*(6), 759–776. https://doi.org/10.1080/00207450290025752.

RStudio Team. (2018). *RStudio: Integrated Development for R (Version 1.2.1335)* [computer software]. RStudio. http://www.rstudio.com.
Salthouse, T. A. (1992). What do adult age differences in the Digit Symbol Substitution Test reflect? *Journal of Gerontology, 47*(3), 121–128. https://doi.org/10.1093/geronj/47.3.P121.
Scheepers, C., & Crocker, M. W. (2004). Constituent order priming from reading to listening: A visual-world study. In M. Carreiras & C. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP, and beyond* (1st. ed.,, pp. 167–185). Psychology Press. https://doi.org/10.4324/9780203509050.
Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49*(3), 238–299. https://doi.org/10.1016/j.cogpsych.2004.03.001.
Sommerfeld, L., Staudte, M., & Kray, J. (2022). Ratings of name agreement and semantic categorization of 247 colored clipart pictures by young German children. *Acta Psychologica, 226*, 103558. https://doi.org/10.1016/j.actpsy.2022.103558.
SR Research (2017). *EyeLink 1000 Plus user manual (Version 1.0.12)*. SR Research.
SR Research. (2019a). *EyeLink Data Viewer (Version 4.1.211)* [computer software]. SR Research. https://www.sr-research.com.
SR Research. (2019b). *EyeLink Experiment Builder (Version 2.2.61)* [computer software]. SR Research. https://www.sr-research.com.
Tallberg, I. M., Carlsson, S., & Lieberman, M. (2011). Children's word fluency strategies. *Scandinavian Journal of Psychology, 52*(1), 35–42. https://doi.org/10.1111/j.1467-9450.2010.00842.x.
Tröger, J., Linz, N., König, A., Robert, P., Alexandersson, J., Peter, J., & Kray, J. (2019). Exploitation vs. exploration—Computational temporal and semantic analysis explains semantic verbal fluency impairment in Alzheimer's disease. *Neuropsychologia, 131*, 53–61. https://doi.org/10.1016/j.neuropsychologia.2019.05.007.
Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*(1), 138–146. https://doi.org/10.1037/0894-4105.11.1.138.
Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73*(2), 89–134. https://doi.org/10.1016/S0010-0277(99)00032-3.
van der Elst, W., Hurks, P., Wassenberg, R., Meijs, C., & Jolles, J. (2011). Animal verbal fluency and design fluency in school-aged children: Effects of age, sex, and mean level of parental education, and regression-based normative data. *Journal of Clinical and Experimental Neuropsychology, 33*(9), 1005–1015. https://doi.org/10.1080/13803395.2011.589509.
van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology, 83*(2), 176–190. https://doi.org/10.1016/j.ijpsycho.2011.09.015.
van Rij, J. (2012). *Pronoun processing: Computational, behavioral, and psychophysiological studies in children and adults*. University of Groningen. University of Groningen Research Output. Doctoral dissertation.
Vergilova, Y., Jachmann, T. K., Mani, N., & Kray, J. (2022). Age-related differences in expectation-based novel word learning. *Psychophysiology, 59*(8). https://doi.org/10.1111/psyp.14030. Article e14030.
Weighall, A. R., & Altmann, G. T. (2011). The role of working memory and contextual constraints in children's processing of relative clauses. *Journal of Child Language, 38*(3), 579–605. https://doi.org/10.1017/S0305000910000267.
Winter, B. (2013). *A very basic tutorial for performing linear mixed effects analyses (Tutorial 2)*. Merced: University of California. https://bodowinter.com/tutorial/bw_LME_tutorial2.pdf.
Winter, B. (2020). Statistics for linguists: An introduction using R. *Routledge*. https://doi.org/10.4324/9781315165547.
Zhang, L., & Knoeferle, P. (2012). Visual context effects on thematic role assignment in children versus adults: Evidence from eye tracking in German. In N. Miyake (Ed.), *Proceedings of the annual meeting of the Cognitive Science Society* (Vol. 34, pp. 2593–2598). https://escholarship.org/uc/item/8hp1f9f7.
Zimmermann, N., de Mattos Pimenta Parente, M. A., Joanette, Y., & Fonseca, R. P. (2014). Unconstrained, phonemic and semantic verbal fluency: Age and education effects, norms and discrepancies. *Psicologia: Reflexão e Crítica, 27*(1), 55–63. https://doi.org/10.1590/S0102-79722014000100007.

Further reading

Bialystok, E., & Poarch, G. (2014). Language experience changes language and cognitive ability. *Zeitschrift Für Erziehungswissenschaft, 17*(3), 433–446. https://doi.org/10.1007/s11618-014-0491-8.
Bobb, S. C., Huettig, F., & Mani, N. (2016). Predicting visual information during sentence processing: Toddlers activate an object's shape before it is mentioned. *Journal of Experimental Child Psychology, 151*, 51–64. https://doi.org/10.1016/j.jecp.2015.11.002.
Brouillette, R. M., Foil, H., Fontenot, S., Correro, A., Allen, R., Martin, C. K., ... Keller, J. N. (2013). Feasibility, reliability, and validity of a smartphone based application for the assessment of cognitive function in the elderly. *PLoS One, 8*(6). https://doi.org/10.1371/journal.pone.0065925. Article e65925.
Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language, 115*(3), 149–161. https://doi.org/10.1016/j.bandl.2010.07.006.
Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology, 39*(2), 133–146. https://doi.org/10.1111/1469-8986.3920133.
Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage.
Friesen, D. C., Luo, L., Luk, G., & Bialystok, E. (2015). Proficiency and control in verbal fluency performance across the lifespan for monolinguals and bilinguals. *Language, Cognition and Neuroscience, 30*(3), 238–250. https://doi.org/10.1080/23273798.2014.918630.
Kidd, C., White, K. S., & Aslin, R. N. (2011). Toddlers use speech disfluencies to predict speakers' referential intentions. *Developmental Science, 14*(4), 925–934. https://doi.org/10.1111/j.1467-7687.2011.01049.x.
Lescht, E. (2017). *Text recognition skills and text recall ability of adults who do and do not stutter in oral and typed modes.* Master's thesis, University of Maryland, College Park. Digital Repository of the University at Maryland.

Winter, R. E., Stoeger, H., & Suggate, S. P. (2021). Fine motor skills and lexical processing in children and adults. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.666200 666200.

Nielsen, T. R., & Waldemar, G. (2016). Effects of literacy on semantic verbal fluency in an immigrant population. *Aging, Neuropsychology, and Cognition, 23*(5), 578–590. https://doi.org/10.1080/13825585.2015.1132668.