Face time: Effects of shyness and attention to faces on early word learning

Matt Hilton
Maastricht University, the Netherlands
Lancaster University, UK

Katherine E. Twomey
University of Manchester, UK

Gert Westermann
Lancaster University, UK

Abstract: Previous research has shown that shyness affects children's attention during the fast-mapping of novel words via disambiguation. The current study examined whether shyness also affects children's attention when eye-gaze cues to novel word meanings are present. 20- to 26-month-old children's (N = 31) gaze was recorded as they viewed videos in which an onscreen actor sat at a table on which one novel and two familiar objects appeared. The actor looked at and labeled one of the objects, using a novel word if the target object was novel. Overall, shyness was associated with a stronger preference for looking at the actor's face, and less time looking at the object being labeled. These effects did not differ when the target object was novel or familiar, suggesting that shyness is related to attentional differences during object labeling generally, rather than specific processes involved in the disambiguation of novel words. No evidence was found of a relation between retention and shyness or attention during labeling.

Keywords: temperament; referent selection; retention; early childhood; eye-tracking

Introduction

During early language development, children are often required to determine word meanings before these meanings can be learned. This first stage of word learning is challenging because of referential uncertainty: There are multiple potential referents for any newly-encountered word, and there exists no reliably unambiguous cue as to the intended referent (Quine, 1960). Children must therefore become skilled at quickly disambiguating the meaning of unfamiliar words, and they make use of a range of different cues to do so. For example, a parent might look in the direction of a bowl containing a pink-colored bobbled fruit that their child has never seen before, and an array of familiar fruits. When the parent then says the novel word “lychee”, their child will typically map this novel word to the unfamiliar fruit, rather than one of the familiar fruits (Halberda, 2006; Mervis & Bertrand, 1994), a behavior known as “fast-mapping” via disambiguation (Carey, 1978; Carey & Bartlett, 1978). Subsequent work has suggested that children can fast-map novel words via disambiguation as early as 17 months of age (Halberda, 2003; but see also Kucker et al., 2018), and by 24 months of age this behavior is reliably demonstrated in lab-based tasks (e.g., Axelsson et al., 2012; Bion et al., 2013; Horst & Samuelson, 2008; Horst et al., 2010). In this example, however, the child need not rely solely on fast-mapping cues to disambiguate the word “lychee”: The child might also be close enough to notice that the parent is looking at the pink-colored fruit. From as early at 3 months of age, babies will shift their attention to look at the target object of somebody else’s gaze (e.g., Hood et al., 1998), and from around 18 months of age, children can map a novel word to an object that is being looked and pointed at (Baldwin, 1993).

Experiments that examine children’s novel word disambiguation typically find variability in children’s performance. For example, in typical tests of fast-mapping via disambiguation, 24-month-old children incorrectly select a familiar object as the referent of a novel word on approximately one quarter of trials (e.g., Axelsson et al., 2012; Horst & Samuelson, 2008). Although children’s incorrect selections on these tasks are often (implicitly) treated as reflective of random error (for example due to a passing distraction) or preference for a particular familiar object (non-linguistic preference effects; e.g., Moore et al., 1999), more recent evidence has indicated that these errors may reflect stable individual differences in children’s novel word disambiguation. Hilton and Westermann (2017) examined whether fast-mapping errors can in part be explained by enduring temperament-based individual differences, rather than in-the-moment random error: They presented 24-month-old children with a typical experimental task and examined whether children’s shyness could explain differences in their novel word disambiguation. Shyness is a biologically-based and enduring temperamental trait characterized as discomfort in (predominantly novel) social situations (Putnam et al., 2006), and it has been shown to affect children’s vocabulary growth (Smith Watts et al., 2014; Spere et al., 2004). Hilton and Westermann found that when presented with an array of one novel and two familiar objects, shyer children were less likely to select the novel object as the referent of the novel word (e.g., in response to the question *where's the koba?*) than less-shy children, indicating that children's fast-mapping via disambiguation is modulated by enduring temperament-related individual differences. Given the face-to-face nature of the task, however, it could not be concluded whether shyer children are generally less likely to disambiguate via fast-mapping, or whether their reduced fast-mapping was reflective of an unwillingness to engage with the task given the discomfort they felt in the novel social situation.

More recent work has attempted to probe the mechanism underlying the relation between shyness and disambiguation. For example, Hilton et al. (2019) removed some of the social demands of the word-learning task by converting it to a looking-while-listening task (e.g., Fernald et al., 1998): 20- to 26-month-old children sat on their parent's lap in a testing booth and viewed images of sets of three objects (one novel, two familiar) on a screen, while a familiar or the novel object was labeled via pre-recorded sentences played through speakers. Even though children were not required to offer a response on this task, shyness modulated their looking patterns across the array of objects during labeling. Specifically, shyness was linked to a reduction in attention to the target object, regardless of whether the heard label was novel or familiar. This finding indicates that shyer children's reduced fast-mapping could be explained in terms of differences in attentional processing. In a version of the same task with 30-month-old children, Axelsson et al. (2022) asked children to point to the referent of the heard word. It was found that children's approachability, a temperamental sub-domain tightly aligned with shyness, was negatively correlated with their pointing accuracy, meaning that shyer children were less likely to correctly select the novel object as the referent of a novel label, specifically on the second trial on which each novel word-object mapping was presented. Interestingly, results of this study also revealed a negative relation between children's temperamental reactivity, defined as children's tendency to feel intense emotions particularly in new contexts, and their looking time to both novel and familiar objects during labeling.

A question arising from this previous work is whether effects of shyness on fast-mapping are specific to disambiguation, or whether they come to bear on children's formation of novel word-object mappings more generally. Given shyer children's specific aversion to unfamiliar people (Putnam et al., 2006), it is plausible that shyness can modulate the formation of novel word-object mappings via social-based cues, such as eye-gaze provided by an unfamiliar adult. Somewhat counterintuitively, shyness is related to greater attention to faces, and in particular to eyes (Brunet et al., 2009; Matsuda et al., 2013; Wieser et al., 2009). These findings have typically been explained in terms of shyer children's hypervigilance to threatening or aversive stimuli. In general, children and adults show more rapid orientation and greater overall attention to threatening or aversive stimuli (e.g., Field, 2006), which is enhanced in shyer individuals who find novel social encounters aversive (e.g., Poole & Schmidt, 2021). In examining the role of shyness in children's formation of novel word-object mappings via eye-gaze cues, it is important to establish whether shyness is related to hypervigilance to faces already during the second and third years of life, the time during which eye-gaze cues play an increasing role in children's word-object mapping (Baldwin, 1993).

It is, however, not immediately clear how a hypervigilance to faces would affect children's novel word mapping via eye-gaze. On the one hand, shyer children's increased attention to faces could support a more rapid and more accurate use of eye-gaze to determine the referent of a spoken word. On the other hand, children's general tendency to look longer at stimuli that they find aversive could mean that shyer children do not follow the eye-gaze cues to the referent, and therefore fail to map the word to the target object. In particular, it would be fruitful to examine these potential effects in the context of novel word disambiguation, given previous interpretations that shyer children's aversion to novel objects (e.g., Kagan et al., 1987; Rothbart, 1988) can explain their reduced target object selection and looking on fast-mapping tasks (Axelsson et al., 2012; Hilton & Westermann, 2017; Hilton et al., 2019). By examining the effect of shyness on looking times to the target object and the face during labeling, we will be better able to understand how aversion to novelty as a marker of shyness comes to bear on division of attention across novel faces and objects.

The formation of novel word-object mappings is, however, only the first stage of word learning. The newly-formed mappings must subsequently be retained, and recent work has argued that attentional processes during the disambiguation of a novel word are critical in determining whether the child will successfully retain the word-object mapping. For example, disambiguation alone is not sufficient to support retention of the newly-formed word-object mapping in 24-month-old children. Instead, retention of the word-object mapping is only demonstrated if the child's attention to the object is heightened following disambiguation (for example, by lifting it up and away from any competitors) while the word is repeated (Horst & Samuelson, 2008). A potential explanation for this finding is that disambiguation is driven by attention to familiar competitors: In order to eventually map the novel word to the novel object, the familiar competitors must first be ruled out as potential referents. Eye-tracking data supports this explanation by showing that looking behavior during disambiguation is characterized by equal looking towards familiar competitors as to the novel object (Hilton et al., 2019; Twomey et al., 2018). Taken together, these findings suggest that attention to familiar competitors is critical for successful fast-mapping, while heightened attention to the target object is critical for retention of the newly formed word-object mapping. Successful word learning is therefore the result of a complex balance of attention across objects during disambiguation. Based on evidence that shyness is also related to a reduced retention of recently-formed novel word-object mappings (Axelsson et al., 2022; Hilton & Westermann, 2017), it is plausible that shyness-related differences in attention distribution during the formation of novel word-object mappings can explain this effect.

The aim of the current study was therefore to examine whether the effects of shyness on children's novel word disambiguation persist when eye-gaze cues to word meaning are also present. 20- to 26-month-old children were tested on an adaption of the looking-while-listening study used by Hilton et al. (2019): Participants were presented with images of one novel and two familiar objects on a screen while their eye movements were measured by an eye-tracker. Critically, the images were accompanied by an onscreen actor who looked at the target object while labeling it. In line with previous work (Axelsson et al., 2012; Hilton et al., 2019; Horst & Samuelson, 2008; Ma et al., 2022; Twomey et al., 2018), we compared children's looking behavior when disambiguating a novel word with their looking behavior when presented with a known word-object mapping, by including trials on which one of the familiar objects was labeled. We then examined whether children's shyness, as measured by the parent-report Early Childhood Behavior Questionnaire (ECBQ; Putnam et al., 2006), was related to their looking across the face and objects during labeling.

Based on previous findings that shyer children are hypervigilant to faces and eyes (e.g., Matsuda et al., 2013), we predicted that shyness would be positively related to looking to the face, and that this increased looking to the face would reduce shyer children's overall looking time to the target object during labeling. However, despite this reduction in looking time to the target object, we speculated that shyer children's hypervigilance to the face could mean that they are more responsive to the eye-gaze cues, and that these cues may serve to focus shyer children's attention to the target object relative to the competitor objects. Given that increased attention to the target object during labeling is related to a greater likelihood of retaining the word-object mapping (Hilton et al., 2019), we also examined whether any effect of shyness on attention to the target during labeling was also related to later retention. If shyer children are more responsive to eye-gaze cues, then any related focus on the target object relative to the competitors could serve to boost retention. Conversely, attention to competitors to rule them out as potential referents is also critical in supporting retention (Halberda, 2006), meaning that it is also possible that any heightened focus on the target object relative to the competitors may serve to weaken retention.

Method

Participants

A total of 31 typically developing children aged 20 and 26 months old took part in the study. All children were typically-developing monolinguals and were from predominantly white, middle-class families living in Lancaster, UK. There were 16 children in the 20-month age group (*M* = 20 m, 11 days; range = 19 m, 19 days to 20 m, 25 days; 7 girls) and 15 children in the 26-month age group (*M* = 26 m, 14 days; range = 25 m, 8 days to 27 m, 8 days; 4 girls). Data from an additional five 20-month-old children were excluded due to equipment error (*n* = 1), or because they were unable to adequately attend to the experiment (e.g., due to distress or refusal to be in the testing suite; $n = 4$), and data from an additional four 26-month-old children were excluded due to equipment error ($n = 1$), because they were unable to adequately attend to the experiment (e.g., due to distress or refusal to be in the testing suite; $n = 2$), or because they showed no looking to the face on every disambiguation trial ($n = 1$). Families were recruited by contacting parents who had previously indicated interest in participating in child development research. Parents' travel expenses were reimbursed, and children were offered a gift of a storybook for participating.

Prior to their visit to the lab parents were requested to complete the Oxford CDI vocabulary checklist (Hamilton et al., 2000) for their children. Some parents could not complete the questionnaire prior to their visit and were therefore asked to take the questionnaire home and mail it back within a week of their visit. Questionnaire data for one 26-month-old child were missing due to the parent not returning the questionnaire and were replaced by the mean. The 20-month-old group had a mean productive vocabulary of 107 words (range = 7-413 words) and a mean receptive vocabulary of 245 words (range = 45-414 words). The 26-month-old group had a mean productive vocabulary of 246 words (range = 58-368 words) and a mean receptive vocabulary of 350 words (range = 232-414 words). As expected, the 26-month-old group had larger receptive and productive vocabularies than the 20-month-old group (receptive: t (29) = 3.32, $p = .002$; productive: t (29) = 3.62, $p = .001$; two-tailed).

Stimuli and design

Each child took part in disambiguation trials, which were presented on a computer screen, and retention trials, which involved the child selecting 3D objects from a tray. Visual stimuli for disambiguation trials consisted of digital photographs of eight objects selected because they are familiar to two-year-old children (ball, boat, car, cup, fork, motorbike, cell phone, shoe) and four novel objects (e.g., a plastic hand massager; see Figure 1). Each picture was of a similar size (approx. 700 x 700 mm) onscreen. Each novel object was assigned one of four novel pseudowords (cheem, koba, sprock, tannin), all of which were plausible English pseudowords and used in previous research (e.g. Hilton et al., 2019). Sixteen randomization orders were created, and each child in both age groups was assigned one of these sixteen orders. Within each order, the objects were randomly grouped into sets of three, with each set consisting of one novel object and two familiar objects. On subsequent retention trials children were presented with 3D objects, in line with previous work examining similar research questions (e.g. Zosh et al., 2013). Stimuli for the warm-up trials consisted of three familiar objects (helicopter, rubber duck, fork), and the novel objects that had been seen during disambiguation were used for the four retention trials. These objects were all of a similar size (approx. 95 x 70 x 50 mm).

A separate video was created for each of the 12 disambiguation trials. Each video began

Figure 1. ***Photographs of objects used in the study. Panel a) shows example familiar stimuli shown on disambiguation trials. Panel b) shows the four novel stimuli.***

by showing an unfamiliar female Caucasian actor in her mid-20s sitting at a table and looking with a neutral expression at the camera. After 600 ms, pictures of three objects then bounced simultaneously onto the table from the bottom of the screen, one to the left, one to the middle, and one to the right (see Figure 2 for an example still) . The actor then looked at the target object (approx. 3000 ms after the trial onset) and labeled it three times in a neutral tone, embedded within a consistent script (Look, it's a _______! Can you see the _______? Wow, it's a _______!). Each sentence was produced and recorded as in real-time, meaning that the precise onset of individual labels varied slightly across trials. After she had finished speaking (approx. 10,400 ms after she began), the actor looked back at the camera with a neutral expression, and the objects disappeared. Each set was presented three times. On the first two presentations, the novel object acted as the target (novel label trials), and on the final presentation a randomly selected familiar object acted as the target (familiar label trials), to ensure that on novel label trials the novel object had not previously been seen on a preceding familiar trial. The order in which the sets were presented was pseudorandomized, with the constraint that no set was presented more than twice successively. Across the three presentations of a given set, the target appeared once on the left, once in the middle and once on the right.

Figure 2. ***Example video still from referent selection trial.***

Procedure

Shyness questionnaire

Parents completed the shyness scale of the Early Childhood Behavior Questionnaire (ECBQ; Putnam et al., 2006) during their visit. In order to reduce demand biases in parents' responses, three other unrelated questions taken from the ECBQ were included within the questionnaire, but these responses were not analyzed. Presenting questions relating to only these two subscales avoided overburdening parents with questionnaires, and such a procedure is in line with previous work using temperament questionnaires (Justice et al., 2008; Rudasill et al., 2014; Spere & Evans, 2009). The ECBQ is a standardized parent report measure of 18- to 36-month-old children's emerging temperament. Twelve items measure the child's shyness, and each item asks parents to rate from 1-7 (1 = never, 7 = always) how often over the past two weeks their child has demonstrated shy-type behaviors (e.g., "when playing with unfamiliar children, how often did your child seem uncomfortable?"). Averaging across the 12 questions (Cronbach's $\alpha = .83$)

yields a score for each child between 1 (not at all shy) and 7 (extremely shy).

Disambiguation trials

Children sat on their parent's lap approximately 60 cm from a computer monitor mounted above a Tobii x120 eye-tracker, which recorded children's gaze data from both eyes at a sampling rate of 60 Hz. Parents were instructed not to speak to their child or look at the screen during stimulus presentation to avoid influencing their child's looking behavior. The experimenter monitored the session via a video camera to ensure that these instructions were complied with. Videos were imported into Tobii Studio (version 3.4) and programmed to run sequentially. Before stimuli were presented, the gaze of each child was calibrated using a five-point procedure: A colorful child-friendly animation (e.g., a wobbling duck) was displayed in the four corners and middle of a 3x3 grid, and calibration accuracy was checked and repeated if necessary. Disambiguation trials followed immediately after calibration.

After every fourth trial, a four-second long child-friendly animation accompanied by an exciting sound effect (e.g., rattling sounds) was displayed in order to keep children's attention on the screen. After disambiguation trials were completed, children took a five-minute break during which they played in an adjacent room. This break was included in line with previous work (e.g., Horst & Samuelson, 2008) to ensure that the subsequent retention phase required recall from long-term memory.

Data coding and cleaning. The raw data files were exported from Tobii Studio (version 3.4) and processed in R (version 4.2.1; R Core Team, 2022) via R Studio (version 1.2.5001; RStudio Team, 2020) with the tidyverse package (version 1.3.2; Wickham et al., 2019). For each participant, the data file showed a timestamp for each data sample and the corresponding x-y coordinates of the child's gaze on the screen. Four square object Areas of Interest (AOIs) were created in the areas of the screen where the stimuli were displayed. All AOIs measured 400 by 400 pixels. An object AOI covered each position in which the objects appeared: left, middle and right. There was a gap of 100 pixels between object AOIs. A margin of 20 pixels separated AOIs from the left and right edge of the screen, and a margin of 40 pixels separated the AOIs from the bottom of the screen. A further AOI covering the position of the actor's face measuring 400 by 400 pixels was defined. AOIs did not overlap. Continuous gaze within an AOI was counted as a fixation. If continuous gaze within an AOI was interrupted for less than 60 ms, this interruption was recoded as a continuation of that fixation, because this was most likely due to blinking or eye-tracking errors rather than the child rapidly re-orienting their attention (0.36 % of data samples were recoded in this way). Only data samples collected following the approximate onset of the target name and until the disappearance of the objects (4500 – 10400 ms from video onset) were analyzed. The proportion of looking time in each AOI was calculated for each trial by dividing the sum of gaze samples in the AOI by the sum of gaze samples that fell into any AOI[1], and these proportions were converted to two log-gaze proportion ratios for analysis (Arai et al., 2007; Borovsky et al., 2016). Proportions of 0 were transformed to 0.01 to allow for log transformation. A face-vs-target log-gaze proportion ratio was calculated by log transforming the proportion looking time to the face divided by proportion looking time to the target object, log(*P*[Face]/*P*[Target]). A log-gaze proportion ratio of zero reflects equal looking across face and target, a positive log-gaze proportion ratio reflects preferential looking to the face, and a negative log-gaze proportion ratio reflects preferential looking to the target object. The magnitude of the log-gaze proportion ratio reflects the strength of the preference. A target-vs-competitors log-gaze proportion ratio was also calculated, log(*P*[Target]/*P*[Competitors]).

Retention Trials

Following the five-minute break, children returned to the testing room to take part in retention trials. Retention trials began with a warm-up task. Children were seated on their parent's lap opposite the experimenter. The experimenter presented the three familiar objects to the child on a tray specially divided into three sections, initially out of reach of the child, for approximately three seconds. Children were then asked to select one of the objects (e.g., *Where's the duck?*), the tray was pushed forward into the child's reach, and their response was recorded. If the child selected the correct object, both the experimenter and parent praised the child, or if the child selected an incorrect object, the experimenter and parent encouraged the child to select the correct one. If the child failed to respond after two further prompts, the tray was removed, and the next trial began. On each subsequent trial, the three objects were rearranged out of sight of the child, and children were asked for a different object. The warm-up task continued until children had selected the correct object three times in a row. Across the warm-up task, each target object appeared in each section of the tray at least once. Retention trials continued in the same manner as the warm-up task, with two differences. First, no praise or encouragement was offered following the child's response: the experimenter simply replied with a neutral *thank you*. Second, retention trials consisted of the novel objects seen during the referent selection phase. On each trial the child was presented with a target alongside two other randomly selected novel objects from the referent selection trials and was asked for the target using the appropriate novel word. There was one retention trial for each novel object; each child therefore participated in four retention trials. The order of retention trials was randomly determined, as was the location of the target on the tray.

[1]e.g., $P[\text{Face}] = \frac{\text{sum samples [Face]}}{\text{sum samples [Face]}+\text{sum samples [Target]}+\text{sum samples [Competitors]}}$

Results

Proportional looking during disambiguation

To determine whether shyness was associated with increased looking to the face relative to the target object, the face-vs-target log-gaze proportion ratios were submitted to a linear mixed effects model (LMEM) with main effects of shyness score (mean-centered across all models) and trial type (sum coded: familiar label trial = 1, novel label trial = -1 across all models), with their interaction, by-participant correlated random slopes for shyness score and intercepts and by-target random intercepts[2]. Log-gaze proportion ratios from six trials were excluded from the analysis because the child looked at neither the face nor the target object. Results are shown in Table 1.

Table 1: Results of linear mixed effects model for face-vs-target log-gaze proportion ratios. Significant predictors are highlighted in bold.

	β	*SE*	*t*	χ^2	*df*	*p*
intercept	0.83	0.22	3.76			
Trial Type	-0.35	0.13	-2.68	5.38	1	.020
Shyness	1.04	0.32	3.28	6.74	1	.009
Trial Type x Shyness	0.14	0.15	0.94	0.87	1	.350

For this analysis, a positive log-gaze proportion ratio indicates preferential looking to the face relative to the target object, and the positive intercept estimate therefore indicates that overall, children looked more to the face than to the target object. Interestingly, the significant negative main effect of trial type indicates that children showed a greater preference to look at the target object relative to the face on familiar label trials (*M* log-gaze proportion ratio = 0.45, *SD* = 1.54) than on novel label trials (*M* = 1.13, *SD* = 1.44). Critically, the significant main effect of shyness indicates that shyness was associated with greater tendency to look at the face, and less looking to the target object. These findings confirm the prediction that shyer children would look more to the face relative to the target object.

Axelsson et al. (2022) found that the relation between approachability and looking behavior differed between the first and second labeling event of the novel object. In the current study, children also saw each novel object labeled on two separate trials.

[2]All linear mixed effects models were conducted using the lme4 package (version 1.1-30 Bates et al., 2015) and initially defined with maximal random effects structures, which were then simplified until convergence (D. J. Barr et al., 2013). *p*-values for fixed effects were obtained using sequential likelihood ratio tests, and *p*-values from follow-up pairwise comparisons were Bonferroni-corrected unless otherwise stated. Initial analyses revealed no effect of age on looking during referent selection, so age was excluded as a fixed factor in models of referent selection to maximize power (cf. Hilton et al., 2019). Estimated random effect variances and R formulae for each model are reported in the supplementary materials.

We therefore ran a subsequent exploratory analysis examining face-vs-target log-gaze proportion ratio on novel label trials only, including fixed factors of shyness and labeling event (sum coded: first labeling event =-1, second labeling event = 1). As expected, analyses confirmed a significant fixed effect of shyness (β = 0.99, *SE* = 0.33, *t* = 3.00, χ^2 = 6.69, *p* = 0.010). The fixed effect of labeling event was marginally non-significant (β = -0.22, *SE* = 0.12, *t* = -1.81, χ^2 = 3.66, *p* = 0.056), suggesting that children tended to look less to the face relative to the target object on the second novel labeling event than the first. Critically, however, no interaction between shyness and labeling event was found (β = 0.19, *SE* = 0.17, *t* = 1.14, χ^2 = 1.29, *p* = 0.26).

The finding that shyer children attended more to the face relative to the target object does not rule out the possibility that these children were more responsive to the eye-gaze cues to disambiguate the heard label. More accurate use of the eye-gaze cues could be reflected by greater attention to the target object relative to the competitor objects. Target-vs-competitor log-gaze proportion ratios were therefore submitted to a LMEM with main effects of shyness score and trial type with their interaction, by-participant correlated random slopes for trial type and intercepts and by-target random intercepts. Log-gaze proportion ratios from 24 trials were excluded from the analysis because the child looked at neither the target object nor the competitor objects. Results can be seen in Table 2.

Table 2: Results of linear mixed effects model for target-vs-competitor log-gaze proportion ratios. Significant predictors are highlighted in bold.

	β	*SE*	*t*	χ^2	*df*	*p*
intercept	0.01	0.17	0.05			
Trial Type	0.34	0.14	2.45	5.49	1	.019
Shyness	-0.44	0.25	-1.80	2.53	1	.110
Trial Type x Shyness	-0.16	0.20	-0.80	0.64	1	.420

For this analysis, a positive log-gaze proportion ratio reflects preferential looking to the target relative to the competitors. The intercept estimate therefore indicates that children overall looked roughly equally across the target object and the competitor objects. The main effect of trial type indicates that children looked preferentially to the target object on familiar label trials (*M* log-gaze proportion ratio = 0.31, *SD* = 1.55) but to the competitors on novel label trials (*M* = -0.30, *SD* = 0.98). These findings suggest that, despite the presence of eye-gaze cues to the target object on all trials, attention to the target was heightened relative to competitors only on familiar label trials. On novel label trials, children still sought to rule out familiar competitors as potential referents. It therefore appears that the eye-gaze cues did not override children's fast-mapping via disambiguation behaviors. We will return to this point in the discussion. Furthermore, no main effect of shyness was found, providing no evidence that shyer children's increased attention to the face served to focus their attention more on the target object relative to the competitors.

We also examined whether the relation between shyness and target-vs-competitor log-gaze proportion ratio differed between the first and second labeling event of the novel object. We therefore ran a further LMEM on data from novel label trials, with main effects of shyness and labeling event (sum coded: first labeling event =-1, second labeling event = 1). These analyses revealed no effect of shyness (β = -0.23, *SE* = 0.32, t = -0.71, χ^2 = 0.61, p = 0.43) and no effect of labeling event (β = -0.10, *SE* = 0.12, t = -0.86, χ^2 = 0.33, p = 0.57). There was also no interaction between shyness and labeling event (β = -0.27, *SE* = 0.18, t = -1.53, χ^2 = 2.30, p = 0.13).

Looking time during disambiguation

The analysis of log-gaze proportion data revealed that shyness modulated children's division of attention across the face and target object during labeling. While this analysis revealed that increased shyness was associated with greater attention to the face relative to the target object, it was unclear whether this effect was related to reduced looking times to the target object. For example, if a highly-attentive child shows a stronger preference for the face relative to the target object, they could still spend longer looking at the target object than a child with a weaker preference for the face and reduced overall looking. A further series of LMEMs were therefore run in order to examine whether differences in children's division of attention, as measured by log-gaze proportion ratios, affected summed looking times to the different AOIs onscreen. Looking time on each trial in seconds was submitted to a LMEM with main effects of shyness score, trial type, and AOI hit type (sum coded: contrast 1: competitor = 1, face = 0, target = -1; contrast 2: competitor = 0, face = 1, target = -1) with their interactions and by-participant random intercepts. Results can be seen in Table 3.

Table 3: Results of linear mixed effects model for looking time during disambiguation. Significant predictors are highlighted in bold.

		β	*SE*	t	χ^2	*df*	p
intercept		1.56	0.06	27.16			
AOI Hit Type	contrast 1	-0.47	0.06	-8.61	211.01	2	< .001
	contrast 2	-0.79	0.06	14.41			
Trial Type		-0.11	0.04	-2.79	7.16	1	.007
Shyness		0.02	0.08	0.02	<0.01	1	.950
AOI Hit Type x Shyness	contrast 1	-0.17	0.08	-2.24	61.88	2	< .001
	contrast 2	-0.56	0.08	7.22			
AOI Hit Type x Trial Type	contrast 1	-0.07	0.06	-1.21	27.33	2	< .001
	contrast 2	-0.21	0.06	-3.81			
Shyness x Trial Type		0.02	0.06	0.40	0.16	1	.691
3-way Interaction	contrast 1	0.09	0.08	1.14	1.29	2	.524
	contrast 2	-0.05	0.08	-0.62	1.29		

The LMEMs revealed a main effect of AOI hit type. As expected based on results of the proportional looking time data, this effect was due to longer looking times to the face (M = 2.41 s; SD = 0.99) than to the target object (M = 1.21 s; SD = 0.68; $\chi^2(1) = 135.29$, $p < .001$) and longer looking time to the face than to the competitor objects (M = 1.16 s; SD = 0.47; $\chi^2(1) = 144.29$, $p < .001$), while looking time to competitors and target did not differ ($\chi^2(1) = 0.48$, $p > .99$).

The main effect of trial type reveals that children looked generally longer to all AOIs on novel label trials (M = 1.67 s; SD = 0.27) than on familiar label trials (M = 1.46; SD = 0.53). There are several potential explanations for this effect. For example, novel label trials were likely more cognitively demanding and therefore required greater attention to determine the correct referent. The experimental design may also explain this effect: The familiar label trials were designed to appear after the child had already seen the novel object labeled twice, so that on familiar label trials participants may have been less attentive due to experimental fatigue.

Critically, an interaction between shyness and AOI hit type was revealed, showing that shyness modulated children's looking times to the three AOI hit types. However, follow-up analysis of simple main effects did not reveal a significant relation between shyness and looking time to the face ($\chi^2(1) = 5.02$, $p = .08$), target ($\chi^2(1) = 5.14$, $p = .07$) or competitors ($\chi^2(1) = 2.87$, $p = .27$). Although not originally planned, we also opted to re-examine these effects using an alternative *p*-value adjustment to better understand which effects might be driving the significant interaction. The Benjamini-Hochberg (Benjamini & Hochberg, 1995) adjustment, which in contrast to the Bonferroni correction attempts to control for the false-discovery rate, indicated that shyness was positively related to looking times to the face ($p = 0.038$), negatively related to looking time to the target ($p = 0.038$), and unrelated to looking time to the competitors ($p = 0.090$). Taken together with the proportional analyses, we therefore tentatively conclude that shyness was associated with a decrease in looking time to the target and an increase in looking time to the face (see Figure 3).

Finally, pairwise comparisons of the interaction between AOI hit type and trial type (see Figure 4) revealed that children looked longer to the face on novel label trials (M = 2.61 s, SD = 1.04) than on familiar label trials (M = 2.00 s, SD = 1.14; $\chi^2(1) = 17.06$, $p < .001$), and they also looked longer to the competitors on novel label trials (M = 1.29 s, SD = 0.58) than on familiar label trials (M = 0.94 s, SD = 0.56; $\chi^2(1) = 11.70$, $p = .002$). These findings indicate that, when the heard label was novel, children's disambiguation was marked by greater looking to the face and the competitor objects. Conversely, children looked longer to the target on familiar label trials (M = 1.42 s; SD = 0.97) than on novel label trials (M = 1.11 s, SD = 0.62; $\chi^2(1) = 9.55$, $p = .006$).

We also examined whether the relation between shyness and looking times to the AOIs differed across first and second labeling events of the novel objects. We therefore ran LMEMs on data from the novel label trials only. Due to convergence-related issues,

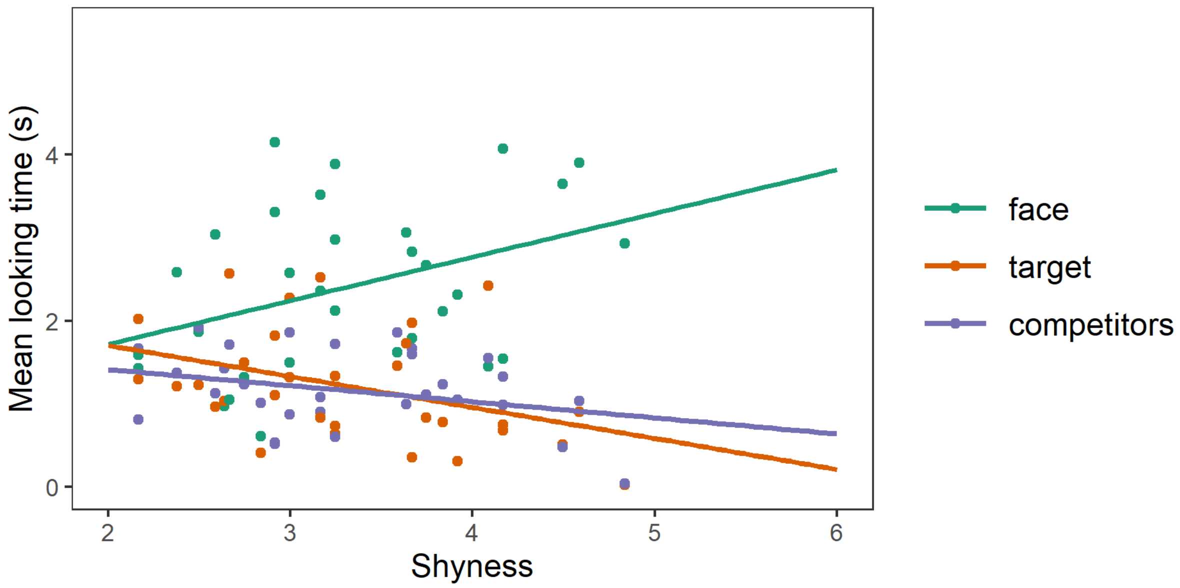

Figure 3. ***Children's shyness scores plotted against their mean looking to each AOI. For illustration, lines are linear regressions.***

three separate LMEMs were run, with summed looking time in each AOI (face, target, competitors) as dependent variable, and shyness, labeling event (sum coded: first labeling event =-1, second labeling event = 1), and their interaction as fixed factors. Full results can be found in the supplementary materials. No significant interaction between shyness and labeling event were found, suggesting that in our study, the relation between shyness and looking time was not modulated by whether the child has already seen the novel object labeled on a previous trial.

Retention Trials

Four children in the 20-month group and one child in the 26-month group did not complete training, and so were excluded from retention analyses. Retention trials were scored 1 if the child selected the correct referent and 0 if they did not. In order to test whether children demonstrated retention above levels expected by chance, a proportion correct retention score was calculated for each child and submitted to a one-sample *t*-test with chance set at 0.33. The 20-month-old group did not demonstrate retention above levels expected by chance alone (M = .38, SD = .27, t (11) = 0.57, p = .578). The 26-month-old group also showed no evidence of retaining the novel label meanings (M = .39, SD = .19, t (13) = 1.24, p = .235). While these analyses reveal that overall children did not retain the label-object associations that were presented during referent selection, in line with Hilton et al. (2019), we next examined whether shyness and proportional target looking during disambiguation predicted retention scores. Trial-by-trial retention scores (correct = 1, incorrect = 0) were submitted to a binomial generalized LMEM with

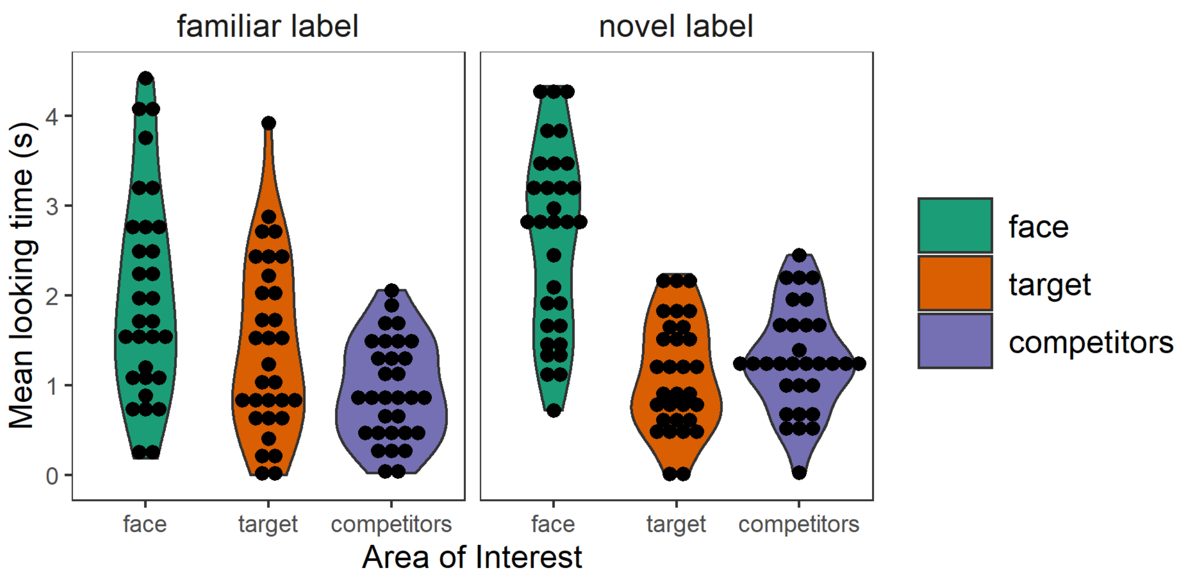

Figure 4. ***Mean looking times in seconds for each participant to each area of interest during disambiguation. The left panel depicts looking on familiar label trials, and the right panel depicts looking on novel label trials.***

main effects of shyness score and target-vs-competitor log-gaze proportion ratio, their interaction, and correlated by-participant random slopes and intercepts for log-gaze proportion ratios and uncorrelated by-target random intercepts and slopes for log-gaze proportion ratios. Results are presented in Table 4, and reveal no significant main effects or interactions, providing no evidence that children's shyness and looking during disambiguation were related to their retention of the label-object associations

Table 4: Results of linear mixed effects model for retention trials.

	β	*SE*	*t*	χ^2	*df*	*p*
intercept	-0.79	0.45	-1.74			
Target/competitor looking[3]	-0.02	0.18	-0.12	0.15	1	.700
Shyness	-0.32	0.33	-0.95	1.48	1	.224
Target/competitor looking x shyness	0.16	0.20	0.81	0.59	1	.444

In order to examine whether looking times to the target or face were significantly related to retention, a further series of LMEMs were run, using summed looking time to the target or face during disambiguation as a main effect alongside shyness score and disambiguation trial type (first novel labeling event vs. second novel labeling event). Full details of model specification and results can be found in the supplementary materials. These analyses also revealed no significant effects of, or interactions with, looking time

[3]Target vs. competitor log-gaze proportion

during disambiguation. These findings therefore provide no evidence that looking behavior during disambiguation is related to retention of the word-object mappings. It is possible that children's poor retention was due to their inability to transfer learning from the screen-based disambiguation task to the live-3D retention task. On the other hand, it is also possible that the presence of the onscreen actor reduced children's overall attention to the target so that they could not sufficiently encode the label-object association. We will return to this point in the discussion.

Discussion

The current work examined whether shyness modulated children's attentional processing during novel word disambiguation, when both eye-gaze and disambiguation cues are provided. The findings suggest that at 20 to 26 months of age, shyness as measured by the ECBQ is related to heightened attention to faces. Critically, this heightened attention to the face did not confer an advantage on shyer children in interpreting eye-gaze cues: shyer children showed the same pattern of looking across target and competitor objects as less-shy children. Instead, the findings indicated that shyer children's heightened attention to the face during labeling reduced their looking time to the target object, which could have weakened their encoding of the word-object mapping. These results could also explain previous findings showing shyer children's reduced novel word disambiguation and retention when measured on a typical face-to-face lab task (Hilton & Westermann, 2017). The current study, however, found no evidence that looking behavior differed with a repeated exposure to the novel word-object mapping, nor that shyness or looking during novel word disambiguation were related to retention of these novel word-object mappings.

The findings that shyer children showed a stronger preference to look at the face, that this preference likely resulted in a decrease in attention to the target object, and that looking to the target object did not differ relative to competitor object looking, indicate that shyer children may struggle in word learning tasks because they do not encode the target object sufficiently to form a robust word-object mapping during referent selection. Previous work has concluded that the formation of a label-object association during referent selection is the product of a complex balance of attention to all available cues. For example, increased attention to a target object during disambiguation has been found to increase the likelihood that this word-object mapping will be retained (Axelsson et al., 2012; Hilton et al., 2019), although removing competitors from the task, despite increasing attention to the target during labeling, reduces retention (Zosh et al., 2013). Similarly, while children make use of eye-gaze cues to form label-object associations from as early as 15 months-of-age (Houston-Price et al., 2006), these cues do not improve learning of label-object associations by 18-month-old children if competitors are highly salient (Moore et al., 1999). The finding that shyness was associated with reduced looking to the target object during labeling replicates those of previous studies that presented images of the target and competitor objects on a blank background (i.e., no onscreen actor or social cues; Axelsson et al., 2022; Hilton et al., 2019). Critically, looking time to the target was reduced not just when the label meaning had to be disambiguated on novel label trials, but also on familiar label trials, when the label meaning was already known. It therefore appears that shyness is related to a general modulation of attentional processes during labeling, instead of individual differences specifically in disambiguation-related cognitive processes.

Our results are in line with previous findings that shyer individuals show heightened attention to eyes and social cues (Brunet et al., 2009; Matsuda et al., 2013). Critically, however, this effect was previously found in older children and in studies that presented stimuli containing only faces. The current study thus extends these previous findings by demonstrating that shyer children also attended preferentially to faces when alternative non-social stimuli (i.e., the target and competitor objects) were displayed. Previous work focusing on adults (Wieser et al., 2009) or older children (Brunet et al., 2009) has suggested that shyer individuals are hyper-vigilant to faces because of heightened self-consciousness (Crozier & Perkins, 2002), meaning that they are more attentive to any social signals that can be conveyed by other people's eyes, and this explanation might apply to the current study. An alternative, lower-level explanation for shyer children's preferential attention to faces in our study could be that it is driven by a formed association between unfamiliar people and feelings of anxiety, because we know that young children show an attentional bias to anxiety-inducing stimuli (e.g., pictures of snakes or angry faces; LoBue & DeLoache, 2010). Critically, shyer children's preferential attention to the face did not result in a more accurate use of gaze cues to disambiguate the novel word: their division of attention across target and competitors on novel label trials did not differ from less-shy children's. This result suggests that increased face looking in shyer children serves to disrupt the attentional processes underlying novel word disambiguation.

Our findings raise some important issues for an understanding of word learning in general. First, despite the presence of eye-gaze cues, all children on average showed the same looking pattern as when objects are displayed on a blank background (Hilton et al., 2019): more looking to the competitor objects than to the target object on novel label trials, and greater attention to the target object than competitor objects only on familiar label trials. Previous work has indicated that even by 18 months of age, eye-gaze cues are not reliably attended to if competitors are highly salient (Moore et al., 1999). However, this same study found that by 24 months of age children will attend more to gaze-cued objects even in the presence of highly salient competitors, which is in contrast to the current study finding no difference in target object looking between 20- and 26-month-old children. Instead, it appears that in our study children did not capitalize on the eye-gaze cues to determine the referent of the novel label, but also disambiguated to eliminate competitors as potential referents. This finding is further evidence of the complex interplay between social and non-social cues to referent selection (e.g., Ma et al., 2022; MacDonald et al., 2017; Yurovsky & Frank, 2017).

Second, we found no relation between looking during disambiguation and children's retention of the new label-object associations. One possibility is that children did not transfer their learning from the 2D pictures of the objects to the actual 3D objects, known as the video deficit effect (Krcmar et al., 2007; Robb et al., 2009; see R. Barr, 2010, for a review), although children have shown no difficulty with this transfer in other studies (e.g., Zosh et al., 2013). Furthermore, despite previous evidence that heightened attention to the target object during disambiguation predicts successful retention of the word-object mapping (e.g., Axelsson et al., 2012), we found no association between looking to the target object and retention. Alongside the video deficit effect, a possible explanation for this finding could be that the presence of the onscreen actor reduced overall attention to the target object during disambiguation below levels sufficient to support retention.

Given that shyness in early childhood is marked specifically by inhibited behavior around unfamiliar adults and in unfamiliar settings (Putnam et al., 2006), it is possible that the shyness-related effects found in the current study are specific to the unfamiliar lab setting combined with the unfamiliarity of the onscreen actor. Follow-up studies could examine this aspect by manipulating the familiarity of the context or the person labeling the objects. The increased availability of mobile and head-mounted eye-tracking would allow, for example, for testing in the child's home or examining children's looking patterns when a familiar adult is labeling the object. These studies would also help us better understand how children's looking behavior on screen-based tasks, such as the current one, relate to children's learning in real-life settings.

Overall, this work shows that shyness exerts a robust effect on attention processing during novel word disambiguation. Specifically, our work demonstrates that the dynamic balance of attention to target object, competitor objects and eye-gaze cues during novel word disambiguation is modulated by shyness. While effects of shyness on social and emotional adjustment have been well-established (e.g., Coplan & Arbeau, 2008), the current study contributes to a growing body of literature that indicates that shyness modulates developing cognitive systems as well. Although shyness in early childhood does not appear to have long-term detrimental direct effects on later language abilities (e.g., Spere & Evans, 2009) or academic achievement (Hughes & Coplan, 2010; Zhang et al., 2017), this growing body of evidence suggests shyness is related to stable individual differences in cognitive processes involved in language development. By better understanding these individual differences, we can begin to support educators and practitioners in determining when children's differential behavior and development are due to normal shyness-related individual differences, or are indicative of more atypical development. Most importantly, work should now begin to further disentangle the dynamic relations between attentional processing, language development and shyness.

References

Arai, M., van Gompel, R. P., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive Psychology*, *54*(3), 218–250. https://doi.org/https://doi.org/10.1016/j.cogpsych.2006.07.001

Axelsson, E., Churchley, K., & Horst, J. (2012). The right thing at the right time: Why ostensive naming facilitates word learning. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00088

Axelsson, E., Othman, N. N., & Kansal, N. (2022). Temperament and children's accuracy and attention during word learning. *Infant Behavior and Development*, *69*, 101771. https://doi.org/https://doi.org/10.1016/j.infbeh.2022.101771

Baldwin, D. A. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology*, *29*, 832–843. https://doi.org/10.1037/0012-1649.29.5.832

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/https://doi.org/10.1016/j.jml.2012.11.001

Barr, R. (2010). Transfer of learning between 2d and 3d sources during infancy: Informing theory and practice [Television and Toddlers: The Message, the Medium, and Their Impact on Early Cognitive Development]. *Developmental Review*, *30*(2), 128–154. https://doi.org/https://doi.org/10.1016/j.dr.2010.03.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bion, R. A., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30months. *Cognition*, *126*(1), 39–53. https://doi.org/https://doi.org/10.1016/j.cognition.2012.08.008

Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016). Semantic structure in vocabulary knowledge interacts with lexical and sentence processing in infancy. *Child Development*, *87*(6), 1893–1908. https://doi.org/https://doi.org/10.1111/cdev.12554

Brunet, P. M., Heisz, J. J., Mondloch, C. J., Shore, D. I., & Schmidt, L. A. (2009). Shyness and face scanning in children. *Journal of Anxiety Disorders*, *23*(7), 909–914. https://doi.org/https://doi.org/10.1016/j.janxdis.2009.05.009

Carey, S. (1978). Linguistic theory and psychological reality. In M. Halle, J. Bresnan, & A. Miller (Eds.). MIT Press.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Proceedings of the Stanford child language conference*, *15*, 17–29.

Coplan, R. J., & Arbeau, K. A. (2008). The stresses of a "brave new world": Shyness and school adjustment in kindergarten. *Journal of Research in Childhood Education*, *22*(4), 377–389. https://doi.org/10.1080/02568540809594634

Crozier, W. R., & Perkins, P. (2002). Shyness as a factor when assessing children. *Educational Psychology in Practice*, *18*(3), 239–244. https://doi.org/10.1080/02667360220000010267

Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, *9*(3), 228–231. https://doi.org/10.1111/1467-9280.00044

Field, A. P. (2006). Is conditioning a useful framework for understanding the development and treatment of phobias? [Anxiety of childhood and adolescence: Challenges and opportunities]. *Clinical Psychology Review*, *26*(7), 857–875. https://doi.org/https://doi.org/10.1016/j.cpr.2005.05.010

Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, *87*(1), B23–B34. https://doi.org/https://doi.org/10.1016/S0010-0277(02)00186-5

Halberda, J. (2006). Is this a dax which i see before me? use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cognitive Psychology*, *53*(4), 310–344. https://doi.org/https://doi.org/10.1016/j.cogpsych.2006.04.003

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a british communicative development inventory. *Journal of Child Language*, *27*(3), 689–705. https://doi.org/10.1017/S0305000900004414

Hilton, M., Twomey, K. E., & Westermann, G. (2019). Taking their eye off the ball: How shyness affects children's attention during word learning. *Journal of Experimental Child Psychology, 183,* 134–145. https://doi.org/https://doi.org/10.1016/j.jecp.2019.01.023

Hilton, M., & Westermann, G. (2017). The effect of shyness on children's formation and retention of novel word–object mappings. *Journal of Child Language, 44*(6), 1394–1412. https://doi.org/10.1017/S030500091600057X

Hood, B. M., Willen, J. D., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science, 9*(2), 131–134. https://doi.org/10.1111/1467-9280.00024

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy, 13*(2), 128–157. https://doi.org/https://doi.org/10.1080/15250000701795598

Horst, J. S., Scott, E. J., & Pollard, J. A. (2010). The role of competition in word learning via referent selection. *Developmental Science, 13*(5), 706–713. https://doi.org/https://doi.org/10.1111/j.1467-7687.2009.00926.x

Houston-Price, C., Plunkett, K., & Duffy, H. (2006). The use of social and salience cues in early word learning. *Journal of Experimental Child Psychology, 95*(1), 27–55. https://doi.org/https://doi.org/10.1016/j.jecp.2006.03.006

Hughes, K., & Coplan, R. J. (2010). Exploring processes linking shyness and academic achievement in childhood. *School Psychology Quarterly, 25*(4), 213–222. https://mu.idm.oclc.org/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=pdh&AN=2010-26288-003&site=ehost-live&scope=site

Justice, L. M., Cottone, E. A., Mashburn, A., & Rimm-Kaufman, S. E. (2008). Relationships between teachers and preschoolers who are at risk: Contribution of children's language skills, temperamentally based attributes, and gender. *Early Education and Development, 19*(4), 600–621. https://doi.org/10.1080/10409280802231021

Kagan, J., Reznick, J. S., & Snidman, N. (1987). The physiology and psychology of behavioral inhibition in children. *Child Development, 58*(6), 1459–1473. Retrieved February 8, 2023, from http://www.jstor.org/stable/1130685

Krcmar, M., Grela, B., & Lin, K. (2007). Can toddlers learn vocabulary from television? an experimental approach. *Media Psychology, 10*(1), 41–63. https://doi.org/10.1080/15213260701300931

Kucker, S. C., McMurray, B., & Samuelson, L. K. (2018). Too much of a good thing: How novelty biases and vocabulary influence known and novel referent selection in 18-month-old children and associative learning models. *Cognitive Science*, *42*(S2), 463–493. https://doi.org/https://doi.org/10.1111/cogs.12610

LoBue, V., & DeLoache, J. S. (2010). Superior detection of threat-relevant stimuli in infancy. *Developmental Science*, *13*(1), 221–228. https://doi.org/https://doi.org/10.1111/j.1467-7687.2009.00872.x

Ma, L., Twomey, K., & Westermann, G. (2022). The impact of perceived emotions on toddlers' word learning. *Child Development*, *93*(5), 1584–1600. https://doi.org/https://doi.org/10.1111/cdev.13799

MacDonald, K., Yurovsky, D., & Frank, M. C. (2017). Social cues modulate the representations underlying cross-situational learning. *Cognitive Psychology*, *94*, 67–84. https://doi.org/https://doi.org/10.1016/j.cogpsych.2017.02.003

Matsuda, Y.-T., Okanoya, K., & Myowa-Yamakoshi, M. (2013). Shyness in early infancy: Approach-avoidance conflicts in temperament and hypersensitivity to eyes during initial gazes to faces. *PLOS ONE*, *8*(6), 1–7. https://doi.org/10.1371/journal.pone.0065476

Mervis, C. B., & Bertrand, J. (1994). Acquisition of the novel name-nameless category (n3c) principle. *Child Development*, *65*(6), 1646–1662. Retrieved October 24, 2022, from http://www.jstor.org/stable/1131285

Moore, C., Angelopoulos, M., & Bennett, P. (1999). Word learning in the context of referential and salience cues. *Developmental Psychology*, *35*(1), 60–68. https://mu.idm.oclc.org/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=pdh&AN=1998-03083-005&site=ehost-live&scope=site

Poole, K. L., & Schmidt, L. A. (2021). Vigilant or avoidant? children's temperamental shyness, patterns of gaze, and physiology during social threat. *Developmental Science*, *24*(6), e13118. https://doi.org/https://doi.org/10.1111/desc.13118

Putnam, S. P., Gartstein, M. A., & Rothbart, M. K. (2006). Measurement of fine-grained aspects of toddler temperament: The early childhood behavior questionnaire. *Infant Behavior and Development*, *29*(3), 386–401. https://doi.org/https://doi.org/10.1016/j.infbeh.2006.01.004

Quine, W. V. O. (1960). *Word and object*. MIT Press.

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Robb, M. B., Richert, R. A., & Wartella, E. A. (2009). Just a talking book? word learning from watching baby videos. *British Journal of Developmental Psychology*, *27*(1), 27–45. https://doi.org/https://doi.org/10.1348/026151008X320156

Rothbart, M. K. (1988). Temperament and the development of inhibited approach. *Child Development*, *59*(5), 1241–1250. Retrieved February 8, 2023, from http://www.jstor.org/stable/1130487

RStudio Team. (2020). *Rstudio: Integrated development environment for r*. RStudio, PBC. Boston, MA. http://www.rstudio.com/

Rudasill, K. M., Prokasky, A., Tu, X., Frohn, S., Sirota, K., & Molfese, V. J. (2014). Parent vs. teacher ratings of children's shyness as predictors of language and attention skills. *Learning and Individual Differences*, *34*, 57–62. https://doi.org/https://doi.org/10.1016/j.lindif.2014.05.008

Smith Watts, A. K., Patel, D., Corley, R. P., Friedman, N. P., Hewitt, J. K., Robinson, J. L., & Rhee, S. H. (2014). Testing alternative hypotheses regarding the association between behavioral inhibition and language development in toddlerhood. *Child Development*, *85*(4), 1569–1585. https://doi.org/https://doi.org/10.1111/cdev.12219

Spere, K. A., & Evans, M. A. (2009). Shyness as a continuous dimension and emergent literacy in young children: Is there a relation? *Infant and Child Development*, *18*(3), 216–237. https://doi.org/https://doi.org/10.1002/icd.621

Spere, K. A., Schmidt, L. A., Theall-Honey, L. A., & Martin-Chang, S. (2004). Expressive and receptive language skills of temperamentally shy preschoolers. *Infant and Child Development*, *13*(2), 123–133. https://doi.org/https://doi.org/10.1002/icd.345

Twomey, K. E., Ma, L., & Westermann, G. (2018). All the right noises: Background variability helps early word learning. *Cognitive Science*, *42*(S2), 413–438. https://doi.org/https://doi.org/10.1111/cogs.12539

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wieser, M. J., Pauli, P., Alpers, G. W., & Mühlberger, A. (2009). Is eye to eye contact really threatening and avoided in social anxiety?—an eye-tracking and psychophysiology study. *Journal of Anxiety Disorders, 23*(1), 93–103. https://doi.org/https://doi.org/10.1016/j.janxdis.2008.04.004

Yurovsky, D., & Frank, M. C. (2017). Beyond naïve cue combination: Salience and social cues in early word learning. *Developmental Science, 20*(2), e12349. https://doi.org/https://doi.org/10.1111/desc.12349

Zhang, L., Eggum-Wilkens, N. D., Eisenberg, N., & Spinrad, T. L. (2017). Children's shyness, peer acceptance, and academic achievement in the early school years. *Merrill-Palmer Quarterly, 63*(4), 458–484. Retrieved February 16, 2023, from http://www.jstor.org/stable/10.13110/merrpalmquar1982.63.4.0458

Zosh, J. M., Brinster, M., & Halberda, J. (2013). Optimal contrast: Competition between two referents improves word learning. *Applied Developmental Science, 17*(1), 20–28. https://doi.org/10.1080/10888691.2013.748420

Data, Code and Materials Availability Statement

Raw data, analysis scripts and digital stimuli used in the current study can be found on the Open Science Framework: https://osf.io/2dhjb/?view_only=05921c9efdc34926a0cbe79ce5508e47. Access to the Early Childhood Behavior Questionnaire (ECBQ) can be requested here: https://research.bowdoin.edu/rothbart-temperament-questionnaires (last retrieved 14/11/2022). The Oxford Communicative Development Inventory (CDI) can be downloaded here: https://www.psy.ox.ac.uk/research/oxford-babylab/research-overview/oxford-cdi (last retrieved 14/11/2022). The Editor agreed an exemption (15/11/2022) to materials-sharing for the ECBQ and Oxford CDI on the basis that both are subject to copyright restrictions.

Ethics statement

Ethical approval for this study was granted by the Lancaster University Research and Ethics Committee. All participants' parents gave informed written consent before taking part in the study.

Authorship and Contributorship Statement

MH wrote the manuscript and collected the data. All authors were involved in study design, analyses, manuscript editing, and contributed intellectually to the manuscript.

Funding Details

This work was supported by the ESRC International Centre for Language and Communicative Development (LuCiD) [ES/L008955/1, ES/S007113/1], an ESRC Future Research Leaders fellowship to KET [ES/N01703X/1] and a British Academy/Leverhulme Trust Senior Research Fellowship to GW [SF150163].

Supplementary Materials: Full model specification and results of retention data

This document outlines the retention-based analyses using looking times (rather than log-gaze proportions) as outcome measure. All fixed effects and interactions were insignificant. The specific results can be found below. As in the main paper, all LMEMs were maximally-defined and then simplified until convergence.

Predicting retention by looking time to the *target* during disambiguation.

Model Specification:

```
retTarSum[[1]] <- glmer(retain ~ sum * tempShy.centered * appearance +
    (1 + sum | ppcode) + (1 + sum | target), data = retainSum_tar,
    family = binomial, control = glmerControl(optimizer = "bobyqa"))
```

Model Result:

Estimate	Std. Error	z value	Pr(>|z|)	factorName	npar	AIC	LRT	p	sig
-1.00000	0.56	-1.79000	0.073	(Intercept)					
0.19000	0.33	0.58000	0.560	sum	1	239.0754	0.52000	0.47022	
0.07400	0.53	0.14000	0.890	tempShy.centered	1	239.6026	1.04900	0.30582	
0.00065	0.31	0.00210	1.000	appearance	1	238.5568	0.00290	0.95687	
-0.35000	0.35	-1.03000	0.300	sum:tempShy.centered	1	244.4936	0.92000	0.33761	
0.03600	0.23	0.16000	0.870	sum:appearance	1	243.6750	0.10000	0.75069	
-0.42000	0.42	-0.99977	0.320	tempShy.centered:appearance	1	243.5744	0.00028	0.98672	
0.40000	0.31	1.27000	0.200	sum:tempShy.centered:appearance	1	245.5741	1.71000	0.19111	

Predicting retention by looking time to the *face* during disambiguation.

Model Specification:

```
retFaceSum[[1]] <- glmer(retain ~ sum * tempShy.centered * appearance +
    (1 | ppcode) + (1 + sum | target), data = retainSum_face,
    family = binomial, control = glmerControl(optimizer = "bobyqa"))
```

Model Result:

Estimate	Std. Error	z value	Pr(>|z|)	factorName	npar	AIC	LRT	p	sig
-0.370	0.62	-0.60	0.55	(Intercept)					
-0.160	0.18	-0.92	0.36	sum	1	235.6462	0.95	0.33049	
-0.974	0.60	-1.61	0.11	tempShy.centered	1	235.2795	0.58	0.44619	
-0.160	0.36	-0.45	0.65	appearance	1	234.9053	0.21	0.64986	
0.290	0.20	1.45	0.15	sum:tempShy.centered	1	240.6485	2.18	0.13941	
0.044	0.13	0.34	0.74	sum:appearance	1	238.5865	0.12	0.72636	
0.200	0.49	0.40	0.69	tempShy.centered:appearance	1	238.5039	0.04	0.84168	
-0.060	0.17	-0.34	0.73	sum:tempShy.centered:appearance	1	240.4640	0.12	0.73311	

Variable descriptions

- **retain:**
  retention score for each novel object (1=retained, 0=not retained)

- **sum:**
summed time looking in target AOI (depending on analysis either face or taregt object, s)
- **tempShy.centered:**
shyness score (centered)
- **appearance:**
looking time on first time the child heard the novel label during labeling, or second time? (-1 = first time, 1 = second time)
- **ppcode:**
participant identifier
- **target:**
novel object identifier

Supplementary Materials: Full model specification and results of looking time analysis with novel labelling event as fixed factor

This document outlines the looking time analyses aimed to examine whether the relation between shyness and looking time to each AOI (face, target, competitors) differed according to the labelling event (i.e., whether it was the first or second trial on which the child saw the target novel object labelled.) Due to failure to converge, it was not possible to include labelling event as a fixed factor alongside AOI. Three separate LMEMs were therefore conducted, with looking time to each of the three AOIs as dependent variable. For each LMEM, shyness, labelling event (sum coded: first labelling event =-1, second labelling event = 1), and their interaction were included as fixed factors. The model specifications and results can be found below

Predicting looking time to the *face* during disambiguation.

Model Specification:

```
sumFACE_model[[1]] <- lmer(sum ~ tempShy.centered * appearance +
    (1 | ppcode) + (0 + tempShy.centered | target), data = sumData.FACE,
    REML = FALSE)
```

Model Result:

Estimate	Std. Error	t value	factorName	npar	AIC	LRT	p	sig
2.63	0.170	15.28	(Intercept)					
0.56	0.250	2.21	tempShy.centered	1	791.71	4.500	0.03400	*
-0.29	0.082	-3.56	appearance	1	798.46	11.250	0.00079	*
-0.17	0.120	-1.45	tempShy.centered:appearance	1	789.21	2.097	0.15000	

Predicting looking time to the *target* during disambiguation.

Model Specification:

```
sumTAR_model[[1]] <- lmer(sum ~ tempShy.centered * appearance +
    (1 | ppcode) + (1 | target), data = sumData.TAR, REML = FALSE)
```

Model Result:

Estimate	Std. Error	t value	factorName	npar	AIC	LRT	p	sig
1.089	0.130	8.37	(Intercept)					
-0.360	0.140	-2.57	tempShy.centered	1	599.47	5.9100	0.015	*
-0.010	0.054	-0.09	appearance	1	593.57	0.0018	0.970	
-0.040	0.076	-0.54	tempShy.centered:appearance	1	595.57	0.2900	0.590	

Predicting looking time to the *competitors* during disambiguation.

Model Specification:

```
sumCOMP_model[[1]] <- lmer(sum ~ tempShy.centered * appearance +
    (1 | ppcode), data = sumData.COMP, REML = FALSE)
```

Model Result:

Estimate	Std. Error	t value	factorName	npar	AIC	LRT	p	sig
1.27	0.098	12.92	(Intercept)					
-0.26	0.140	-1.88	tempShy.centered	1	639.77	3.430	0.064	
-0.02	0.061	-0.34	appearance	1	636.60	0.250	0.620	
0.15	0.086	1.74	tempShy.centered:appearance	1	638.34	3.018	0.082	

Variable descriptions

- **retain:**
  retention score for each novel object (1=retained, 0=not retained)
- **sum:**
  summed time looking in target AOI (depending on analysis either face or taregt object, s)
- **tempShy.centered:**
  shyness score (centered)
- **appearance:**
  looking time on first time the child heard the novel label during labeling, or second time? (-1 = first time, 1 = second time)
- **ppcode:**
  participant identifier
- **target:**
  novel object identifier