Infants Show Early Comprehension of Basic Colour Words

Samuel H. Forbes* and Kim Plunkett
University of Oxford

Abstract

Previous research has highlighted the difficulty that infants have in learning to use colour words. Even after acquiring the words themselves, infants are reported to use them incorrectly, or over-extend their usage. We tested 146 infants from 5 different age groups on their knowledge of 6 basic colour words, *red*, *green*, *yellow*, *blue*, *black* and *white*, using an inter-modal preferential looking task. The results showed that infants show reliable comprehension of colour words as early as 19-months of age. No order of acquisition effects were observed. In addition, infants' behaviour in the task was facilitated by the provision of redundant noun information, "Look at the red car", and even general referential NPs, "Look at the red one", with greater looking to the target than when the colour label was not presented in adjective position, "Look, red". The findings indicate that colour words may be learned with greater ease than previously thought, verifying recent parental reports showing similar findings. The findings also suggest that 19-month olds have already developed an expectation that colour labels should occur in adjectival position.

Keywords: Eye-tracking, Colour words, Word learning

Introduction

The nature and timing of colour word learning has been a topic of much debate. Early reports suggested that colour words were produced correctly as late as 7 years of age (Heider, 1971). Later evidence suggested more precocious knowledge of colour terms before 4 years (Bornstein, 1985; Franklin, 2006; Pitchford &

*Contact author. Email: samuel.forbes@psy.ox.ac.uk. SHF is funded by the Rhodes Trust. The authors declare no conflict of interest. The authors are grateful to Ioana Duta for assistance with calculating the stimuli lengths.

Mullen, 2002), and some level of production during the second year of life (Mervis, Bertrand, & Pani, 1995; Shatz, Behrend, Gelman, & Ebeling, 1996). Even when the colour labels themselves are acquired, their usage was thought to be riddled with errors. Children regularly have more difficulty with some colours than others, particularly non-focal colours (Andrick & Tager-Flusberg, 1986; O'Hanlon & Roberson, 2006; Pitchford & Mullen, 2001, 2005), and having learned them they apply the colour terms inconsistently (e.g. Kowalski & Zimiles, 2006; Pitchford & Mullen, 2003; Rice, 1980; Roberson, Davidoff, Davies, & Shapiro, 2004; Sandhofer & Smith, 1999; Soja, 1994).

Studies to date have therefore found it difficult to establish a clear timeline for when colour words are learned, which has lead to various theories about why they might be so difficult for children to learn. Explanations have varied from children's inability to abstract the category boundaries in order to map the categorical colour words (Andrick & Tager-Flusberg, 1986), to infants lacking a conceptual representation of colour (Kowalski & Zimiles, 2006), as well as linguistic and attentional constraints (O'Hanlon & Roberson, 2006).

More recent work has found evidence that colour words may be learned like slow-mapped categories, with a partial comprehension preceding production, and that comprehension is slowly refined as the infants learn more about the category (Wagner, Dobkins, & Barner, 2013; Wagner, Jergens, & Barner, 2014, 2018). These claims find further support in studies of cross-linguistic parental report data (Forbes & Plunkett, in press), which found that in 11 different languages, parents reported an early colour word comprehension prior to production. One of the aims of the present study is to examine the acquisition of colour words further with experimental, behavioural data.

Forbes and Plunkett (in press) reported evidence for colour word comprehension beginning much earlier than previously found, with around 50% of infants comprehending the four basic colours by 21 months of age. In an investigation of the relationship between colour word comprehension and production, Wagner et al. (2018) found signs that infants with a mean of 23 months of age, comprehended colour terms, based on evidence from eye-tracking experiments and parental report. Yet despite evidence verifying parental report as a reliable estimation of children's word learning (e.g. Dale, 1991; Mills, Coffey-Corina, & Neville, 1993, 1997), there has also been debate about the validity of parental reports (Houston-Price, Mather, & Sakkalou, 2007; Tomasello & Mervis, 1994). The Forbes and Plunkett (in press) findings were also in stark contrast to previously collected behavioural data (e.g. Pitchford & Mullen, 2002; Sandhofer & Smith, 1999), making it unclear whether their findings are a result of the methods used, or consistent with children's real comprehension of colour words; a question conflated by the fact that many previous behavioural studies used production as a measure, or required the child to interact with the experimenter.

Measuring colour word comprehension with a behavioural task is complicated as it raises the question of colour preferences. In a controlled trial, if an infant reaches for the red shape having been prompted to "*find the red one*," that may be as much due to red being a colour of interest as to their possible comprehension of the word "*red*." Colour preferences in infants have been well-documented, finding that even in pre-linguistic infants, infants look longer at red hues, as opposed to green hues (Franklin, Bevis, Ling, & Hurlbert, 2010), and that the preference for red is consistent across context (Franklin, Gibbons, Chittenden, Alvarez, & Taylor, 2012). Despite these findings, behavioural measures have yet to evaluate colour word comprehension uncontaminated by colour preferences, one of the key aims of this study.

In controlled experimental conditions, infants often have great difficulty mapping adjectives to object properties (Mintz & Gleitman, 2002; Waxman & Markow, 1995). Colour is no exception to this. Children finding mapping a novel adjective on to a colour to be a very difficult task (e.g. Booth & Waxman, 2009). Additional linguistic context may make the mapping process easier, such that a child may find it easier to attend to an object property when a specific noun is provided. For example, the child may affix their gaze on the red car more readily when hearing, "*look at the red car*," than when hearing "*look at the red one*" (Mintz & Gleitman, 2002). The present study also aims to address this question, by manipulating the context in which the target colour words are presented, and examining whether the context influences their recognition of the target.

The present study addresses each of the above questions by measuring colour word comprehension using Intermodal Preferential Looking (IPL) procedures (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987). Each IPL trial can be examined in two steps: the pre-naming phase, which provides information about the baseline preferences infants have for one colour over another; and the post-naming phase, which measures their responses to auditory prompts. Based on the parental report studies of Forbes and Plunkett (in press), it was hypothesised that the four chromatic colour words would be learned by the 24 month-old mark, earlier than previous behavioural experiments have shown. In addition, the present study aimed to examine how the provision of different types of sentential information affects infants' comprehension of a colour word, by using three structures that differentially highlight the adjectival status of the colour word. We predicted that the infants would look more reliably to the target when the colour word was embedded in a prototypical adjectival position, in line with the findings of Mintz and Gleitman (2002).

Methods

Participants

Participants were recruited in 5 age groups: 30 participants at 1;0 were recruited for a baseline no-comprehension control, as they were unlikely to understand colour words at that age. 28 participants ranging from 3;0 – 4;0 upwards were also recruited as the comprehension control group, as they were likely to all comprehend the colour words by that age. In between these groups, 29 participants at 1;4, 31 participants at 1;9, and 28 participants at 2;0 were recruited as the main experimental groups, for a total of $N = 146$ participants. An additional 23 participants were excluded for fussiness or parental interference with the task, while an additional 5 participants were excluded for failing to complete at least one trial with each colour as both distractor and target. Participant information can be found in Table 1.

Table 1
Descriptive statistics for participants included in study.

Age Group	*N*	Mean Age (months)	SD (months)
12	30	11.84	0.70
16	29	15.96	0.70
19	31	19.69	0.73
24	28	24.30	0.36
48	28	53.46	18.78

All participants were contacted after recruitment at the local maternity ward or online. Participants with one parent or grandparent with colour vision problems were not tested for this study. All participants were monolingual, learning English as their first language.

Materials

Auditory stimuli were recorded by a native female speaker of Southern British English (SBE), speaking slowly and clearly in an infant-directed manner. The auditory stimuli consisted of three different sentence types: sparse (*"Look, red!"*), general (*"Look at the red one!"*), and informative (*"Look at the red car!"*). Note that in all cases, attention to the colour label alone is sufficient to succeed in identifying the target. Both the colour and the named object varied depending on what was shown on the screen.

Visual stimuli were all objects that should be familiar to infants in daily life, such as vehicles, items of clothing, or furniture. Each object was chosen to be an object without a typical colour, and that could be easily recoloured. In each trial, the same object was presented on both the left and the right of the screen, varying only in the colour. Objects could be any one of six colours: *red*, *blue*, *green*, *yellow*, *black*, or *white*, and each colour was selected to be a typical example of the colour category, and confirmed to be so by independent observers as well as during pilot testing. Where necessary, objects were recoloured in the GNU Image Manipulation Program (GIMP, www.gimp.org).

Participants saw each colour three times as a target, one corresponding to each of the sentence types, for a total of 18 trials. Participants were randomly assigned to different lists, in order to counter-balance which target colours appeared against which distractor colours, as well as counterbalancing which colours appeared with which objects. Trials were left-right randomised. All trials were run using a custom script in MATLAB, and recorded using a Tobii TX300 eye-tracker, recording at 120Hz.

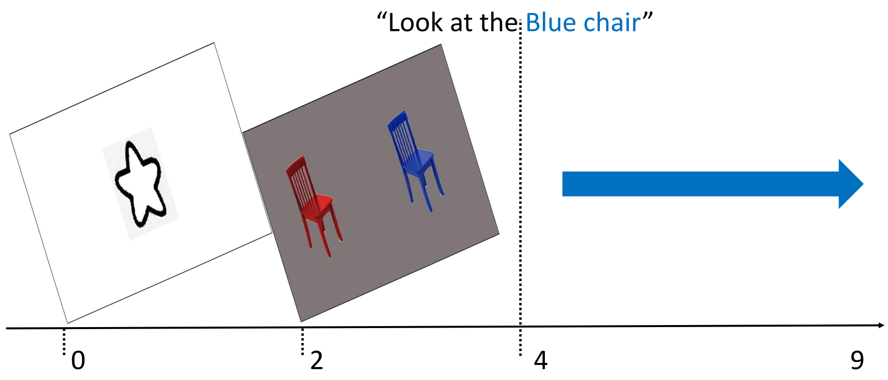

Figure 1. Time course of a typical trial.

Procedure

On arrival at the lab, participants and caregivers were shown to a playroom to allow the participant to familiarize themselves with the laboratory settings. The study was approved by the University of Oxford Medical Sciences Interdivisional Research Ethics Committee, reference number: MS-IDREC- C1-2015- 071 (project title: Adjective and object property comprehension in children aged 3-36 months). During this time, caregivers were asked to fill out consent forms, as well as a parental report, asking whether their child comprehends, or comprehends and says each of the 11 basic colour words (for details see Forbes & Plunkett, in press). After this warming-up period, participants were seated on the lap of the caregiver, roughly 75cm from the eye-tracker and presentation screen.

The experiment commenced with a nine-point calibration sequence, which was repeated until at least 7 of the nine points were calibrated successfully, after which the trials commenced. Each trial lasted for nine seconds, the first two seconds of which was an attractive attention getter designed to orient the participant's attention to the centre of the screen. Immediately after, the two images appeared on the screen, on a neutral grey background. The auditory stimuli were presented so that the onset of the target colour word occurred exactly 2 seconds after the images appeared. Trials continued for another 5 seconds after the onset of the target colour word. For each participant, target/distractor colour pairings were counterbalanced, e.g., if a participant saw a blue chair as a target against a red chair as a distractor (Figure 1), then they would also see the red chair as a target against the blue chair as a distractor. Comparing each target colour to multiple distractor colours in this way dramatically reduces the likelihood of participants using mutual exclusivity to search for the named target, meaning that they are most likely to have to rely on their knowledge of the target colour label.

Analysis

Data for infant fixations were extracted with a custom MATLAB script. A fixation was defined as a stable gaze in one location for at least 100ms, allowing for a small amount of dispersion to account for the unsteadiness of the infant gaze. The area of interest around each image was expanded slightly to allow for the same unsteadiness, so that the borders of each image were expanded by 25%. Trials were removed prior to analysis if more than 60% of the trial was lost due to the infant focussing attention away from the screen. This threshold was set more generously than usual to account for the fact that infants are more likely to lose attention in such long trials. For each analysis, the variable of colour was dummy coded, while age and time elapsed during the trial were treated as continuous numeric variables.

Analysis was completed in R, using the MASS package (Ripley et al., 2017) and eyetrackingR (Dink & Ferguson, 2015). In the pre-naming phase, the data were aggregated across the pre-naming period so that for each participant, each colour could be compared to every other colour.

For the post-naming phase, the decision was made prior to analysis to use data from 0-3000ms after the target word onset, as we hypothesised that as the trial proceeded, the influence of colour preference may overrule the effect of naming. In the post-naming phase, there were two main analyses. In the first, a naming score was calculated for each participant. The naming score was the proportion of looking to the named colour in the first 3000ms after target word onset minus the proportion of looking to that colour before a target was named.

In the second post-naming analysis, data were analysed using a binomial mixed-effects mode. For this analysis, rather than modelling the proportion of looks to the target, which would allow the colour preferences of each infant to bias the result (as each infant may have individual preferences), data were aggregated to obtain the number of looks to each colour in each time bin for each participant when that colour was the *named target* and aggregated again for each colour when it was the *distractor*. The proportion we examine in the post-naming phase is that of the proportion of looks to any given colour when it was named versus when it was not, for each participant, colour, and time bin. The proportion calculated is thus the number of looks to the target colour when it is the named target, divided by when that same colour is the distractor, for each participant and time bin. In other words, for a given participant i and colour j:

$$Proportion_{ij} = \frac{Target_{ij}}{Target_{ij} + Distractor_{ij}}$$

In the sentence type analysis, data were aggregated to calculate the proportion of looks to the named target colour for each participant and sentence type.

Results

Pre-naming Phase

In the pre-naming phase, the data can be used to analyse the overall baseline colour preferences of the participants; the purpose of which is to determine whether baseline colour preferences will affect infant looking to the named target. The proportions of looking to each colour against each other colour can be seen in Figure 2. The figure is a matrix of preference for each colour against each other colour, where red suggests a preference for looking to that colour, and yellow suggests a preference for looking away from that colour. The figure indicates a strong preference to look to red over most colours, and a strong preference to look to any other colour, when the colour shown is white.

The pre-naming phase data were fitted with a multilevel linear regression using the package lme4 (Bates et al., 2017). The model included Colour 1 and Colour 2 as fixed effects (see Figure 2), and varied the intercept for each participant, to allow for individual variance. The model coefficients (Table 2) reinforce the pattern depicted graphically in Figure 2, demonstrating strong evidence for looking toward red, and for looking away from white. While these are only compared to black in the model, they reinforce the pattern that can seen in Figure 2.

Table 2
Model coefficients for pre-naming phase. Colours are compared to black. Results for the second colour are identical but reversed, due to the nature of the data.

	Estimate	Std. Err	df	t value	Pr(>|t|)
(Intercept)	0.50	0.02	2484	26.09	<0.001
Colour1Blue	-0.02	0.02	2484	-1.11	0.269
Colour1Green	0.01	0.02	2484	0.47	0.637
Colour1Red	0.04	0.02	2484	2.26	0.024
Colour1White	-0.08	0.02	2484	-4.13	<0.001
Colour1Yellow	-0.02	0.02	2484	-0.86	0.389

These findings highlight the need to correct for colour preference in the analysis of the post-naming data, as infants show a strong preference for red, and a

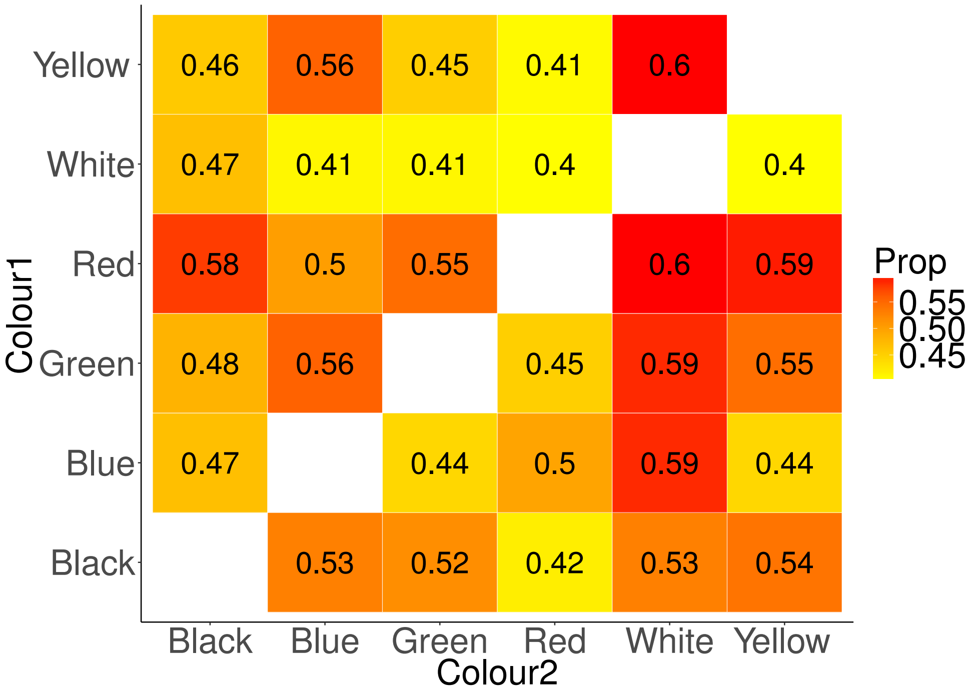

Figure 2. The proportions of looking to any target colour (Colour 1) over any other colour (Colour 2). Red indicates looking to that colour above 50%, yellow indicates looking below 50%. Proportions are listed in each box.

strong preference for any colour over white. The proportions also suggest some basic evidence for a preference for green over blue and yellow, for black over blue, green, and yellow, and for yellow over blue. The results of this analysis are consistent with previous reports of infants preferring red hues over other hues, but in contrast to previous work we do not find a strong preference for blue hues (e.g. Franklin et al., 2010; Teller, Civan, & Bronson-Castain, 2004).

Post-naming Phase

In the post-naming phase, looking was first aggregated across the first 3000ms of trial time to calculate whether colour word responses improved as the trial time increased. Participants were assessed on their looking to the target after the colour word was named, compared to before the target was named; consistent looking to the named target when prompted would suggest comprehension of the target colour word. This naming score at each age group was compared to the null hypothesis of no difference between the two time periods ($\mu = 0$) with One Sample t-tests. At 12 months, there was no evidence that the infants comprehended colour words ($t(29) = 0.245$, $p = 0.808$, 95%CI = -0.021 – 0.027), nor was there any evidence of colour word comprehension at 16 months ($t(28) = 0.482$, $p = 0.634$, 95%CI = -0.016 – 0.026). However, for the subsequent three age groups, there

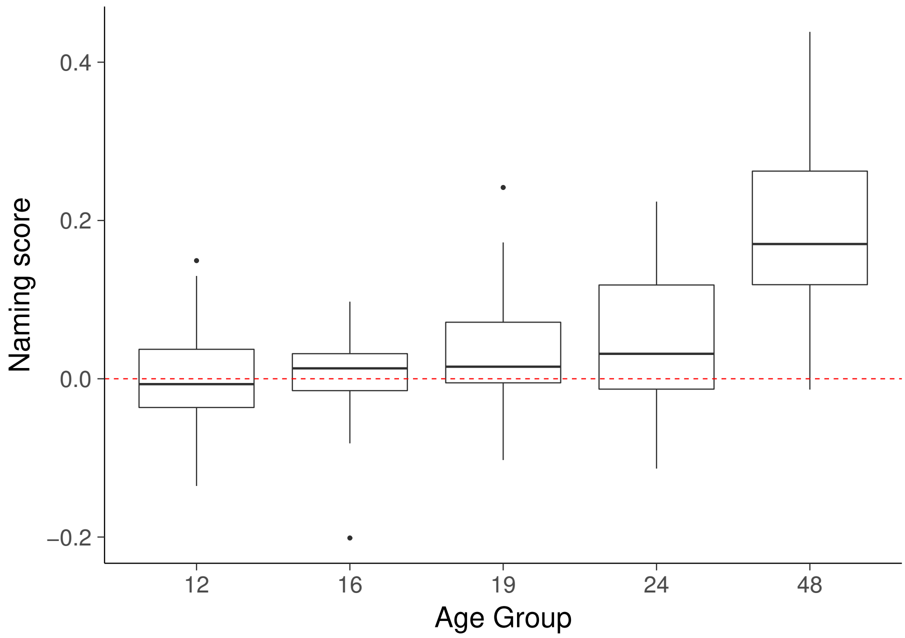

Figure 3. Naming score in each age group. The naming score represents the proportion looking to the target colour after it is named, minus the average looking proportion to that target colour in the pre-naming phase.

was strong evidence that participants comprehended colour words, looking consistently to the target after it was named: at 19 months ($t(30) = 3.029$, $p = 0.005$, 95%CI = 0.012 – 0.064), 24 months ($t(27) = 2.309$, $p = 0.029$, 95%CI = 0.004 – 0.074), and at 48 months ($t(27) = 8.639$, $p < 0.001$, 95%CI = 0.149 – 0.242). Overall looking proportions in each age group can be viewed in Figure 3.

In order to further analyse the differences in colours and looking over the trial time, looks to each colour when it was the *named target* and when it was the *distractor* were modelled with a binomial logistic mixed-effects regression, using the function glmmPQL in R. The regression was fitted with quartic orthogonal polynomials of the time elapsed after target word onset (Mirman, 2014). The numeric variable of participant age and the categorical variable of colour were included in the model. In addition, both the intercept and the slope of colour were allowed to vary for each participant, in order to allow for the fact that comprehension of different colours may vary greatly between individuals. The full list of effects can be viewed in Table A1.

The regression analysis demonstrated strong evidence for an effect of both the linear time term and the cubic time term, as well as for an interaction between both of those time terms and the age of the participant. There was no strong evidence for an effect of colour, although there was evidence for interaction effects between some of the time terms and red and blue, suggesting that overall looking proportions for each of the colours may not have varied much, although there were

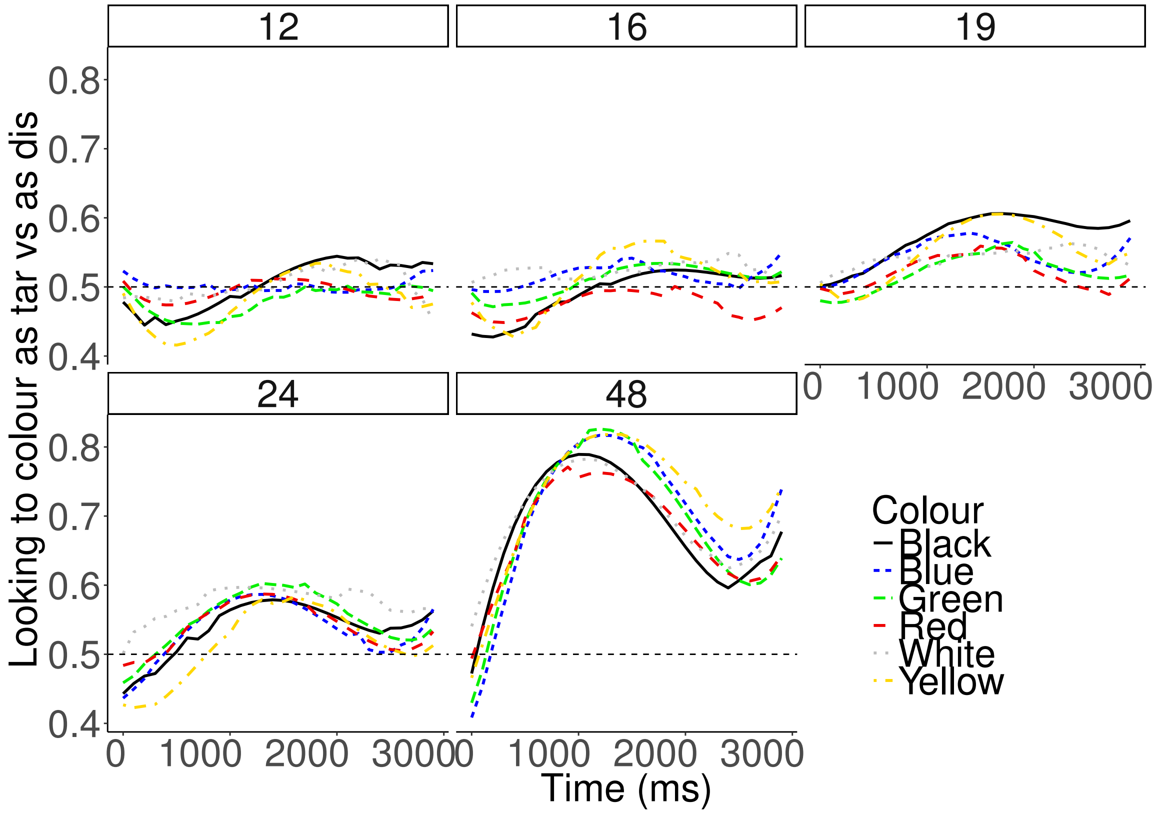

Figure 4. Looking proportion of looks to the named colour when it is the distractor vs looks to the same colour when it is the distractor. 0ms is when the target is named.

some differences in the looking patterns for each colour.

The model fit (Figure 4) shows looking proportions for all colours at around chance at both 12 and 16 months, then at 19 months there is consistent looking to the target above chance, which becomes slightly more consistent at 24 months. The 48 month-olds consistently look to the target. The model fit also demonstrates very little difference between the colours, with looking to the target largely at chance for all six colours at 16 months, and above chance for the age groups thereafter.

Comparison to parental reports

Participants' performance in the eye-tracking task was compared with parental reports of the participants' understanding of these colour terms. The parental report data were derived from the reports they were asked to fill out when arriving at the lab (see the Supplementary Materials as well as Forbes & Plunkett, in press, for more information on the parental report). Participants were marked as comprehending the colour words in the eye-tracking task if their adjusted target versus distractor proportion for each colour (used in the above model) across all trials for that colour exceeded 0.55.

Collapsing across age, a Chi-squared test was performed on whether or not they looked more to the colour word when it was the target than when it was the distractor, versus whether they were judged to have comprehended the colour term according to their parents. A strong association was found between parental report data and eye-tracking data for the colour word comprehension ($\chi^2(1) = 44.207$, $p < 0.001$).

Participants were then taken as "knowing" colour words in general if they knew four or more colour words according to each measurement (Parental Report Comprehends, Parental Report Comprehends and Produces, Eye-tracking Comprehends). The proportions of participants who knew colour words according to each measure was compared in each age group with a binomial regression, treating Age as a categorical variable to allow for analysis of the difference in each age group. A comparison can be seen in Figure 5.

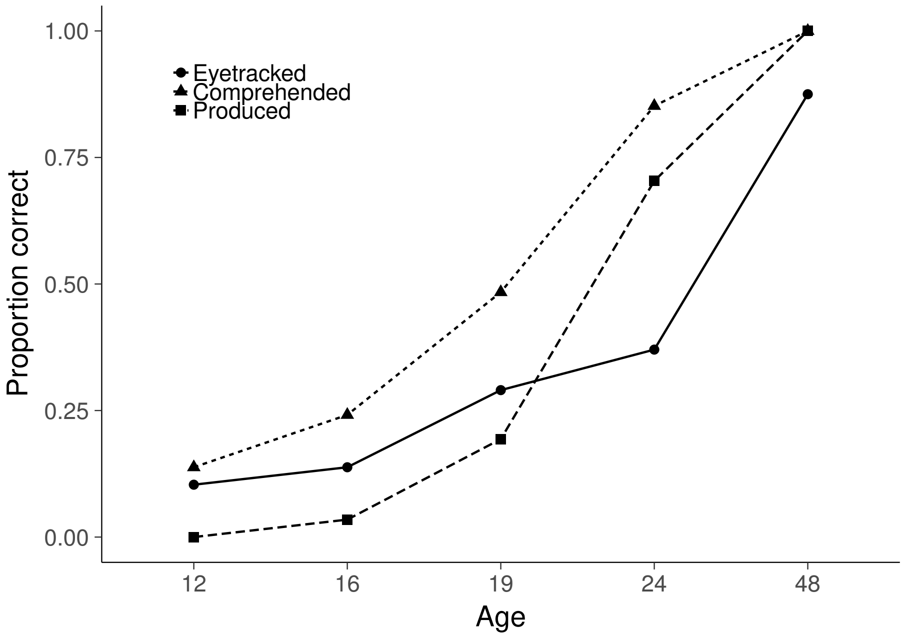

Figure 5. Proportions of participants who were judged to know four or more colour words according to three measures: marked as comprehending on the parental report, marked as comprehending and producing on the parental report, or based on their Eye-tracking results.

The results suggest a clear effect of Age group, and a possible interaction with comprehends as measured on the parental report at age 24 months. Figure 5 demonstrates the close relationship between the eye-tracking results and the parental report results, where the eye-tracking results follow a similar trajectory to the parental report results. The full results can be seen in Table A2.

Sentence type analysis

There were three types of utterance used to introduce the colour terms:

1. Sparse: *"look, red!"*

2. General: *"look at the red one!"*

3. Informative: *"look at the red chair!"*

The effect of sentence structure on participant performance was analysed using a binomial mixed-effects regression, with quartic polynomials as above. The age of the participants and the sentence type were included as fixed effects, and the intercept and slope of Stimulus type were allowed to vary for each participant. Lengths of each stimulus from target word onset until auditory stimulus offset are recorded in Table 3.

Table 3
Mean lengths and standard deviations for each stimulus type following target word onset.

Stimulus	Mean	SD
1	0.716	0.068
2	0.901	0.034
3	1.200	0.151

The model coefficients (Table A3) suggest strong evidence for a difference between the stimulus types (especially the difference between *sparse* and *informative*, $t = 2.883$, $p = 0.004$), and also in their interactions with all of the polynomial time terms. There was also strong evidence for an interaction between the differences in stimulus type and age ($p = 0.004$ and $p = 0.002$, respectively). This demonstrates that the stimuli types that gave participants more contextual information better enabled them to locate the target.

Figure 6 demonstrates the model fit. Of particular interest is the 16m age group, where the model fit shows that having more context (the *general* and *informative* cases respectively) enables the infants to look to the target above chance, although this is not true in the *sparse* condition. This demonstrates how being in the prototypical adjectival position may enhance infants' comprehension of colour words.

Discussion

This study has demonstrated several important aspects of children's colour word learning. First, the study has found strong evidence for early colour word learning in British infants, supporting the recent parental report analysis by

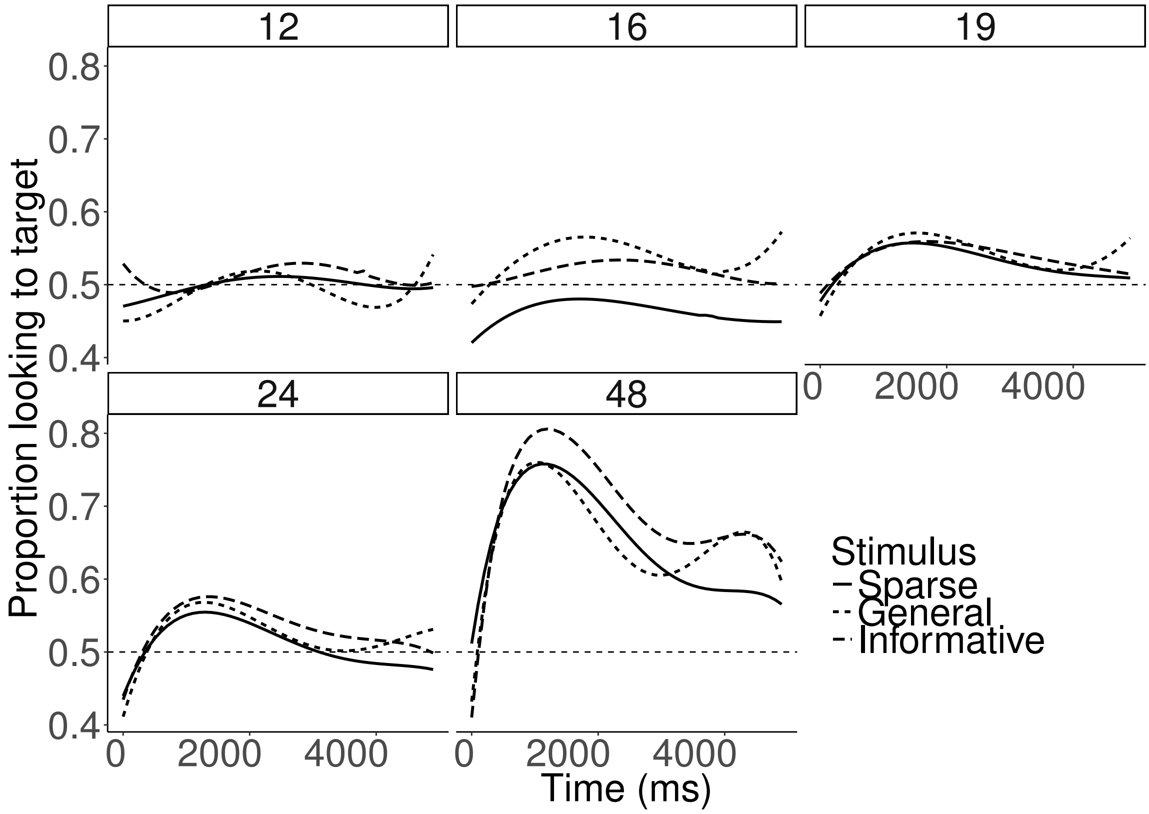

Figure 6. Looking proportion to the target for different sentence types.

Forbes and Plunkett (in press). The present study finds evidence of comprehension of basic colour word knowledge as early as 19 months, much earlier than was found in many previous behavioural analyses (e.g. Pitchford & Mullen, 2002; Sandhofer & Smith, 1999). In fact, Forbes and Plunkett's (in press) parental report study found that colour words were only comprehended by around 25% of infants at 19 months; in contrast the present study found reliable looking to the target at that age across all 6 colours tested, suggesting that British parents may be very conservative when estimating the comprehension of abstract word categories of their children, a possibility previously suggested by Styles and Plunkett (2009). The present study also builds on that data by demonstrating the validity of parental report in evaluating children's knowledge of abstract categories of words such as colour words, supporting the findings of Wagner et al. (2018).

Our study also highlights the role that sentence structure plays in revealing children's understanding of the meaning of colour words. In addition to reinforcing previous findings (Mintz & Gleitman, 2002), this raises important considerations about the role that a colour word plays in the context of a sentence. The colour word being in the penultimate position may confirm to toddlers that it is the property (of colour in this case), and not the object label that is being addressed – suggesting the possibility of an early understanding of how object properties should modify objects. While Mintz and Gleitman employed a very different paradigm to investigate the nature of adjectives, they found that when the target was named, rather than just using "one," subjects were more likely to extract the property information from the linguistic signal.

The present findings suggest that infants expect colour words to describe a noun, rather than be the object of a sentence itself. As colour words, and adjectives in general, primarily describe the properties of an object, they appear to be most informative and easiest to decipher when more context is provided about the object which they describe. Mintz and Gleitman (2002) infer from their results that labelling the noun gives parameters to understand the adjective, an explanation that may have some weight in this circumstance given the real-world stimuli employed in the present study. The results are consistent with an account that supposes that colour words are processed more effectively when given in context, being the difference between comprehending and not comprehending at 16 months.

Finally, this study demonstrates strong preferences for red objects over other colours, demonstrating the importance of controlling for colour preferences. While this will likely be affected to an extent by the choice of background colour, it lends further support to previous studies that have found a preference for the colour red (Franklin et al., 2010).

There are some key differences found between the present study and our recent parental report study (Forbes & Plunkett, in press). One such example is that very little difference was found between the six colours, whereas the parental report study found that black and white were learned after the four basic chromatic colour words. An explanation for this may be that the gap between each of the age groups in the current study is sufficiently large that the differences in when the colour words are learned are not apparent. Forbes and Plunkett (in press) found that the gap between the colours was at most a few months, suggesting it may not be apparent in the present study, where the gap between age groups is 3-4 months at least. In addition, Forbes and Plunkett find that there is a possible slight advantage to learning blue over other colours in British English, again not reflected in the present study, due to the size of the age gap between each group. The results of the present study also support the results of (Wagner et al., 2018), who showed early comprehension in infants around 23 months of age, and possibly as young as 18 months. In contrast to Wagner et al., the present study also examines the effect of age on colour word comprehension, demonstrating when early comprehension begins.

It is important to note that the present study only utilises six basic colour words, and only typical examples of each of these terms. While our results suggest that comprehension occurs much earlier than previously thought, the findings may reflect an extremely basic comprehension of typical examples, not an adult-like understanding of the colour word. Early comprehension may reflect only the beginning of a gradual, slow-mapped process (Wagner et al., 2013, 2014, 2018), after which the infant's comprehension will subsequently become more adult-like with each interaction, until their understanding extends to the boundaries of each colour word. An area of interest for further studies would be to demonstrate how this process occurs, through longitudinal analysis.

Infant word learning can occur in a number of ways. Studies have shown that infants have a propensity to fast-map words under certain conditions (e.g. Heibeck & Markman, 1987), but in general, word learning is a slow-mapped procedure over time (McMurray, Horst, & Samuelson, 2012). In this sense, colour word learning is not unique, with early comprehension preceding production, but occurs slowly over time as infants determine the location of the boundaries of each colour word. It is possible that while a great deal of attention has been given to the difficulties that infants have with learning colour words, it is simply the case that their errors are more obvious than those involving concrete nouns (Yurovsky, Wagner, Barner, & Frank, 2015).

The present study demonstrates that British infants begin the process of comprehending colour words as early as 19 months, and slowly start refining their comprehension over time with age. While there is little doubt that there are manifold reasons that infants may struggle with the mapping of colour words onto the continuous spectrum of colour (Franklin, 2006; Kowalski & Zimiles, 2006; O'Hanlon & Roberson, 2006), they are still able to learn colour words with great efficiency, in much the same way as they do for other classes of words.

References

Andrick, G. R., & Tager-Flusberg, H. (1986). The acquisition of colour terms. *Journal of Child Language*, *13*(1), 119–134. doi: 10.1017/S0305000900000337

Bates, D. M., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., ... Eigen, C. (2017). Fitting linear mixed-effects models using lme4. *Journal of statistical software*, *67*(1), 1 – 113. doi: 10.18637/jss.v067.i01

Booth, A. E., & Waxman, S. R. (2009). A horse of a different color: Specifying with precision infants' mappings of novel nouns and adjectives. *Child Development*, *80*(1), 15–22.

Bornstein, M. H. (1985). On the development of color naming in young children: Data and theory. *Brain and Language*, *26*(1), 72–93. doi: 10.1016/0093-934X(85)90029-X

Dale, P. S. (1991). The validity of a parent report measure of vocabulary and syntax at 24 months. *Journal of speech and hearing research*, *34*(3), 565–571. doi: 10.1016/0165-5876(92)90087-6

Dink, J. W., & Ferguson, B. (2015). *eyetrackingR: An R Library for Eye-tracking Data Analysis.*

Forbes, S. H., & Plunkett, K. (in press). Linguistic and cultural variation in early color word learning. *Child Development*.

Franklin, A. (2006). Constraints on children's color term acquisition. *Journal of Experimental Child Psychology*, *94*(4), 322–327. doi: 10.1016/j.jecp.2006.02.003

Franklin, A., Bevis, L., Ling, Y., & Hurlbert, A. (2010). Biological components of colour preference in infancy. *Developmental Science*, *13*(2), 346–354. doi: 10.1111/j.1467-7687.2009.00884.x

Franklin, A., Gibbons, E., Chittenden, K., Alvarez, J., & Taylor, C. (2012). Infant color preference for red is not selectively context specific. *Emotion*, *12*(5), 1155–1160. doi: 10.1037/a0025333

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14*(01), 23. doi: 10.1017/S030500090001271X

Heibeck, T. H., & Markman, E. M. (1987, August). Word learning in children: an examination of fast mapping. *Child development*, *58*(4), 1021–1034.

Heider, E. R. (1971). *"Focal" color areas and the development of color names.* (Vol. 4) (No. 3). US: American Psychological Association. doi: 10.1037/h0030955

Houston-Price, C., Mather, E., & Sakkalou, E. (2007). Discrepancy between parental reports of infants' receptive vocabulary and infants' behaviour in a preferential looking task. *Journal of Child Language*, *34*(4), 701–724. doi: 10.1017/S0305000907008124

Kowalski, K., & Zimiles, H. (2006). The relation between children's conceptual functioning with color and color term acquisition. *Journal of Experimental Child Psychology*, *94*(4), 301–321. doi: 10.1016/j.jecp.2005.12.001

McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Selection and Slow Associative Learning. *Psychological review*, *119*(4), 831–877. doi: 10.1037/a0029872.Word

Mervis, C. B., Bertrand, J., & Pani, J. R. (1995). Transaction of cognitive-linguistic abilities and adult input: A case study of the acquisition of colour terms and colour-based subordinate object categories. *British Journal of Developmental Psychology*, *13*(3), 285–302. doi: 10.1111/j.2044-835X.1995.tb00680.x

Mills, D. L., Coffey-Corina, S. A., & Neville, H. J. (1993). Language Acquisition and Cerebral Specialization in 20-Month-Old Infants. *Journal of Cognitive Neuroscience*, *5*(3), 317–334. doi: 10.1162/jocn.1993.5.3.317

Mills, D. L., Coffey-Corina, S. A., & Neville, H. J. (1997). Language comprehension and cerebral specialization from 13 to 20 months. *Developmental Neuropsychology*, *13*(3), 397–445. doi: 10.1080/87565649709540685

Mintz, T. H., & Gleitman, L. R. (2002). Adjectives really do modify nouns: The incremental and restricted nature of early adjective acquisition. *Cognition*, *84*(3), 267–293. doi: 10.1016/S0010-0277(02)00047-1

Mirman, D. (2014). *Growth Curve Analysis and Visualization Using R Analysis and Visualization Using R*. Boca Raton, FL: Chapman & Hall / CRC Press.

O'Hanlon, C. G., & Roberson, D. (2006). Learning in context: Linguistic and attentional constraints on children's color term learning. *Journal of Experimental Child Psychology*, *94*(4), 275–300. doi: 10.1016/j.jecp.2005.11.007

Pitchford, N., & Mullen, K. (2001). Conceptualization of perceptual attributes: a special case for color? *Journal of Experimental Child Psychology*, *80*, 289–314. doi: 10.1006/jecp.2001.2634

Pitchford, N., & Mullen, K. (2002). Is the acquisition of basic-colour terms in young children constrained? *Perception*, *31*(11), 1349–1370. doi: 10.1068/p3405

Pitchford, N., & Mullen, K. (2003). The development of conceptual colour categories in pre-school children: Influence of perceptual categorization. *Visual Cognition*, *10*(1), 51–77. doi: 10.1080/713756669

Pitchford, N., & Mullen, K. (2005). The role of perception, language, and preference in the developmental acquisition of basic color terms. *Journal of Experimental Child Psychology*, *90*(4), 275–302. doi: 10.1016/j.jecp.2004.12.005

Rice, M. (1980). *Cognition to language categories: Categories, word meaning, and training*. Baltimore MD: University Park Press.

Ripley, B. D., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., & Firth, D. (2017). *Package ' MASS '*.

Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. R. (2004). The development of color categories in two languages: a longitudinal study. *Journal of experimental psychology. General*, *133*(4), 554–71. doi: 10.1037/0096-3445.133.4.554

Sandhofer, C. M., & Smith, L. B. (1999). Learning color words involves learning a system of mappings. *Developmental Psychology*, *35*(3), 668–679.

Shatz, M., Behrend, D., Gelman, S. A., & Ebeling, K. S. (1996). Colour term knowledge in two-year-olds: evidence for early competence. *Journal of Child Language*, *23*(1), 177–199. doi: 10.1017/S030500090001014X

Soja, N. N. (1994). Young children's concept of color and its relation to the acquisition of color words. *Child Development*, *65*(3), 918–937.

Styles, S., & Plunkett, K. (2009). How do infants build a semantic system? *Language and Cognition*, *1*(1), 1–24. doi: 10.1515/LANGCOG.2009.001

Teller, D. Y., Civan, A., & Bronson-Castain, K. (2004). Infants' spontaneous color preferences are not due to adult-like brightness variations. *Visual neuroscience*, *21*(3), 397–401. doi: 10.1017/S0952523804213360

Tomasello, M., & Mervis, C. B. (1994). The instrument is great but measuring comprehension is still a problem. *Monographs of the Society for Research in Child Development*, *59*(5), 174–179. doi: 10.1111/j.1540-5834.1994.tb00186.x

Wagner, K., Dobkins, K., & Barner, D. (2013). Slow mapping: Color word learning as a gradual inductive process. *Cognition*, *127*(3), 307–317. doi: 10.1016/j.cognition.2013.01.010

Wagner, K., Jergens, J., & Barner, D. (2014). Partial Color Word Comprehension

Precedes Production. *Proceedings CogSci*, 1724–1729.

Wagner, K., Jergens, J., & Barner, D. (2018). Partial color word comprehension precedes production. *Language Learning and Development*, 1–21. doi: 10.1080/15475441.2018.1445531

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: evidence from 12- to 13-month-old infants. *Cognitive psychology*, *29*(3), 257–302. doi: 10.1006/cogp.1995.1016

Yurovsky, D., Wagner, K., Barner, D., & Frank, M. C. (2015). Signatures of Domain-General Categorization Mechanisms in Color Word Learning. *Proceedings CogSci*.

Appendix

Table A1

Full list of effects for logistic mixed-effects regression on the post-naming data. "ot" refers to the orthogonal time terms. Colours are compared to black, values are estimated in log-odds.

	Value	Std.Error	DF	*t*-value	*p*-value
(Intercept)	1.226	1.946	25131	0.630	0.529
ot1	-915.378	194.255	25131	-4.712	<0.001
ot2	512.535	664.883	25131	0.771	0.441
ot3	-566.541	103.023	25131	-5.499	<0.001
ot4	153.179	241.368	25131	0.635	0.526
Age	0.220	0.909	144	0.242	0.809
ColourBlue	-0.630	2.702	25131	-0.233	0.816
ColourGreen	-0.339	2.687	25131	-0.126	0.900
ColourRed	0.803	2.661	25131	0.302	0.763
ColourWhite	-4.525	2.766	25131	-1.636	0.102
ColourYellow	1.336	2.764	25131	0.483	0.629
ot1:Age	646.923	90.756	25131	7.128	<0.001
ot2:Age	3.144	310.536	25131	0.010	0.992
ot3:Age	370.680	48.103	25131	7.706	<0.001
ot4:Age	50.328	112.735	25131	0.446	0.655
ot1:ColourBlue	419.477	270.791	25131	1.549	0.121
ot1:ColourGreen	10.708	268.538	25131	0.040	0.968
ot1:ColourRed	446.353	265.935	25131	1.678	0.093
ot1:ColourWhite	120.443	276.990	25131	0.435	0.664
ot1:ColourYellow	-202.788	276.776	25131	-0.733	0.464
ot2:ColourBlue	-236.221	924.158	25131	-0.256	0.798
ot2:ColourGreen	-129.556	919.177	25131	-0.141	0.888
ot2:ColourRed	305.124	910.192	25131	0.335	0.737
ot2:ColourWhite	-1554.771	945.458	25131	-1.644	0.100
ot2:ColourYellow	537.384	945.263	25131	0.569	0.570
ot3:ColourBlue	312.015	143.905	25131	2.168	0.030
ot3:ColourGreen	60.671	142.695	25131	0.425	0.671
ot3:ColourRed	313.148	141.338	25131	2.216	0.027
ot3:ColourWhite	114.232	147.156	25131	0.776	0.438
ot3:ColourYellow	-55.933	147.234	25131	-0.380	0.704
ot4:ColourBlue	-137.079	335.924	25131	-0.408	0.683
ot4:ColourGreen	-104.083	334.257	25131	-0.311	0.756
ot4:ColourRed	108.681	330.963	25131	0.328	0.743

ot4:ColourWhite	-575.569	343.558	25131	-1.675	0.094
ot4:ColourYellow	202.749	343.961	25131	0.589	0.556
Age:ColourBlue	0.941	1.259	25131	0.747	0.455
Age:ColourGreen	0.268	1.255	25131	0.214	0.831
Age:ColourRed	-0.007	1.238	25131	-0.006	0.996
Age:ColourWhite	1.289	1.277	25131	1.010	0.313
Age:ColourYellow	0.349	1.280	25131	0.272	0.785
ot1:Age:ColourBlue	-116.466	126.186	25131	-0.923	0.356
ot1:Age:ColourGreen	-67.787	125.037	25131	-0.542	0.588
ot1:Age:ColourRed	-285.077	123.930	25131	-2.300	0.021
ot1:Age:ColourWhite	-132.281	127.462	25131	-1.038	0.299
ot1:Age:ColourYellow	-47.353	128.453	25131	-0.369	0.712
ot2:Age:ColourBlue	332.262	430.750	25131	0.771	0.441
ot2:Age:ColourGreen	109.984	429.409	25131	0.256	0.798
ot2:Age:ColourRed	-1.086	423.693	25131	-0.003	0.998
ot2:Age:ColourWhite	426.918	436.488	25131	0.978	0.328
ot2:Age:ColourYellow	89.917	437.829	25131	0.205	0.837
ot3:Age:ColourBlue	-99.832	67.069	25131	-1.489	0.137
ot3:Age:ColourGreen	-56.833	66.443	25131	-0.855	0.392
ot3:Age:ColourRed	-176.324	65.916	25131	-2.675	0.007
ot3:Age:ColourWhite	-84.713	67.725	25131	-1.251	0.211
ot3:Age:ColourYellow	-55.821	68.412	25131	-0.816	0.415
ot4:Age:ColourBlue	144.359	156.679	25131	0.921	0.357
ot4:Age:ColourGreen	72.962	156.288	25131	0.467	0.641
ot4:Age:ColourRed	-0.810	154.151	25131	-0.005	0.996
ot4:Age:ColourWhite	146.084	158.712	25131	0.920	0.357
ot4:Age:ColourYellow	39.653	159.518	25131	0.249	0.804

Table A2
Results of the logistic model run comparing parental report to eye-tracking results.

	Estimate	SE	z value	Pr(>|z|)
(Intercept)	-2.16	0.61	-3.542	<0.001
Age16	0.33	0.81	0.402	0.688
Age19	1.27	0.73	1.741	0.082
Age24	1.63	0.73	2.236	0.025
Age48	4.11	0.87	4.732	<0.001
variableComprehended	0.33	0.81	0.402	0.688
variableProduced	-24.94	86550.00	0	1.000
Age16:variableComprehended	0.36	1.07	0.338	0.736
Age19:variableComprehended	0.50	0.97	0.516	0.606
Age24:variableComprehended	1.95	1.06	1.85	0.064
Age48:variableComprehended	24.65	86930.00	0	1.000
Age16:variableProduced	23.45	86550.00	0	1.000
Age19:variableProduced	24.41	86550.00	0	1.000
Age24:variableProduced	26.34	86550.00	0	1.000
Age48:variableProduced	49.92	122700.00	0	1.000

Table A3

Model coefficients of model fit comparing the type of sentence. Stimulus Type refers to the sentence type, and are compared to sentence type 1.

	Value	Std.Error	DF	t-value	p-value
(Intercept)	0.229	0.271	21716	0.845	0.398
ot1	-7.039	3.095	21716	-2.275	0.023
ot2	7.088	2.722	21716	2.603	0.009
ot3	-4.474	1.708	21716	-2.620	0.009
ot4	1.373	0.732	21716	1.876	0.061
Age	-0.370	0.124	144	-2.987	0.003
Stimulus2	0.781	0.375	21716	2.084	0.037
Stimulus3	1.077	0.374	21716	2.883	0.004
ot1:Age	7.364	1.421	21716	5.182	<0.001
ot2:Age	-7.155	1.252	21716	-5.715	<0.001
ot3:Age	4.095	0.786	21716	5.208	<0.001
ot4:Age	-1.125	0.335	21716	-3.357	0.001
ot1:Stimulus2	-8.431	4.398	21716	-1.917	0.055
ot1:Stimulus3	-12.793	4.385	21716	-2.918	0.004
ot2:Stimulus2	7.271	3.868	21716	1.880	0.060
ot2:Stimulus3	10.631	3.858	21716	2.756	0.006
ot3:Stimulus2	-5.044	2.426	21716	-2.079	0.038
ot3:Stimulus3	-6.220	2.420	21716	-2.570	0.010
ot4:Stimulus2	2.880	1.038	21716	2.774	0.006
ot4:Stimulus3	1.915	1.034	21716	1.851	0.064
Age:Stimulus2	-0.498	0.172	21716	-2.898	0.004
Age:Stimulus3	-0.549	0.173	21716	-3.173	0.002
ot1:Age:Stimulus2	6.087	2.022	21716	3.010	0.003
ot1:Age:Stimulus3	7.342	2.035	21716	3.607	<0.001
ot2:Age:Stimulus2	-4.968	1.781	21716	-2.790	0.005
ot2:Age:Stimulus3	-5.999	1.794	21716	-3.344	0.001
ot3:Age:Stimulus2	3.422	1.118	21716	3.062	0.002
ot3:Age:Stimulus3	3.458	1.126	21716	3.070	0.002
ot4:Age:Stimulus2	-1.653	0.475	21716	-3.481	<0.001
ot4:Age:Stimulus3	-1.098	0.478	21716	-2.299	0.022