BUCLD 50 Proceedings
To be published in 2026 by Cascadilla Press
Rights forms signed by all authors

What Modulates Toddlers' Use of Familiar Words in Resolving Novel Noun Reference? A Preferential Looking Study

Lean Luo, Xiaolu Yang, Stella Christie, and Rushen Shi

1. Introduction

The words a noun co-occurs with can provide useful information for listeners to navigate among its potential referents, which is a crucial step in novel noun learning. However, it remains unclear to what extent young children can exploit such information. This is in part because existing evidence suggests that they do not always make effective use of co-occurring words to resolve a noun's reference, even when all words involved are familiar to them. One such prominent case documented in the literature is their failure to recruit the information provided by noun-modifying adjectives to resolve reference in an adult-like way (e.g., Fernald et al., 2010; Ninio, 2004; Thorpe et al., 2006). For instance, Fernald et al. (2010) found that when 30-month-old English-acquiring children heard *Can you find the blue car?* they did not orient to the target referent more quickly in the context where the adjective was informative (e.g., a blue car and a red house were presented) than in the context where it was not (e.g., the context contained a blue car and a blue house). By contrast, 36-month-olds showed a clearer facilitative effect when they heard an informative prenominal adjective.

This finding raises a concern with respect to young children's capacity for semantic integration more generally. As discussed in Fernald et al. (2010), one possible interpretation is that only older children can integrate the meanings of the elements in a modified noun phrase efficiently (see also Davies et al., 2021 for further integration evidence in three-year-olds). Relatedly, Redolfi & Melloni (2025) showed in an eye-tracking study that Italian-learning children (2;4-5;3) were significantly slower and less accurate than adults in shifting their gaze to the target upon hearing the adjective critical for the identification task (e.g., *orso grande* 'big bear' in a visual context with both a big bear and a small bear). However, in their study, children's integration performance was better for some

*Lean Luo, Tsinghua University and University of California, Davis, leluo@ucdavis.edu; Xiaolu Yang, Tsinghua University, xlyang@mail.tsinghua.edu.cn; Stella Christie, Tsinghua University, christie@mail.tsinghua.edu.cn; Rushen Shi, Université du Québec à Montréal, shi.rushen@uqam.ca. This research was supported by the National Social Science Grant of China (21BYY019) to Xiaolu Yang. We are grateful to all members of the Language Acquisition Lab at Tsinghua University for the feedback and help with data collection. We also thank all participants in the study.

adjectives (e.g., *black*) than the others (e.g., *big*). In other words, integration seems to be selectively more difficult for certain words. In this light, an apparent integration failure need not stem (solely) from the difficulty with the integration process *per se*, but can also be due to children's less efficient or less accurate representations of the words to be integrated.

Our study compares 30-month-old Mandarin-learning toddlers' ability to use different adjectives and numerals in resolving novel noun reference. The primary goal is to further examine how earlier findings on young children's ability to integrate noun phrase elements generalize to a novel noun identification setting when the integration process involves different items. Processing adjective- and numeral-containing noun phrases poses a similar integration challenge, in that both require restricting a set of entities with a certain (either adjectival or cardinality) property. This is also reflected in theoretical studies that assign numerals and adjectives the same semantic type, and by extension, same computation (e.g., set intersection) underlying their combination with nouns (e.g., Rothstein, 2013, though see Kennedy, 2015 for a different approach). Further, as in many other languages, adjectives and numerals share a similar surface position in Mandarin when they occur pre-nominally and form a compositional relation with nouns. As shown in (1), combinations of nouns with numerals require classifiers (CL), and those with adjectives are typically mediated by the functor DE.

(1) a. yi ge dangao
one CL cake
'one cake'
b. piaoliang de dangao
pretty DE cake
'pretty cake'

Even though adjectives and numerals may invoke the same integration process when they combine with nouns, the representations needed to build prior to integration can differ substantially. As already alluded to in previous studies reporting children's integration difficulties with adjectives, the semantics of adjectives can be context-dependent (or noun-dependent as discussed in Thorpe et al., 2006, p.631). For instance, the word *small* receives a very different interpretation when the context is about elephants as opposed to mice (the same also applies to color adjectives like *black*; see Medin & Shoben, 1988). Therefore, constructing their semantic representations often requires the establishment of a comparison class, as achieved by the *for*-PP in *small for an elephant* via explicitly restricting the *domain* of the adjective; and this manifests most clearly in relative gradable adjectives, the special class that includes items such as *big* and *small* (see Kennedy, 2007). By contrast, the interpretation of numerals is typically not context-dependent, thus obviating the need for a comparison class or domain restriction.

Taken together, these similarities and differences provide a basis for further

probing the nature of children's integration failures documented in the literature. If the integration process *per se* is vulnerable for young children, such failures should extend to numerals as well. Otherwise, failures should selectively occur, contingent upon the difficulty of building the semantic representations to be integrated.

2. The present study

In the current study, we asked whether children's integration performance was uniformly weak at 30 months, or it was weak only when the lexical representations to be integrated were challenging to build. We adopted the Intermodal Preferential Looking Paradigm (IPLP; see Hirsh-Pasek & Golinkoff, 1996), a method widely used in studies testing young toddlers. As in previous studies, children were presented with both audio and visual stimuli during the experiment, with their looking behavior recorded for later analysis.

2.1. Participants

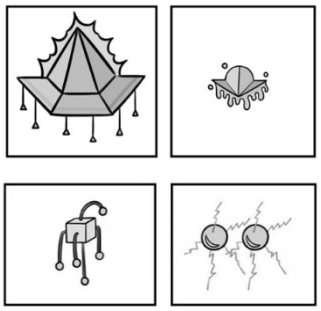
We tested twenty-four 30-month-olds acquiring Mandarin Chinese as their native language (mean age: 2;6;21; 13 males and 11 females). All of them were recruited in Beijing China, reported to be typically-developing by their guardians and receiving minimal to no dialect exposure. Written informed consents were obtained from their guardians before the experiment began. Another three toddlers participated but their data were excluded from the final analysis because they were not able to finish the experiment.

2.2. Stimuli

The critical items in the audio stimuli were adjective- and numeral-containing noun phrases. We used novel items as head nouns to avoid any pre-existing occurrence preferences that might affect the integration difficulty of the noun phrases. Specifically, we chose six disyllabic novel words from Miao (2023), all of which were formed by putting together two existing morphemes in Mandarin. These nouns were then combined with one pre-nominal element, namely an adjective or a numeral, to create the noun phrases whose integration would be tested. As in (1), all adjectives were followed by the functor DE, and numerals by the classifier *ge* in the noun phrases—both mediating elements encode very little semantic information. For the adjective items, we used their reduplicated forms that are known to serve a descriptive function rather than a restrictive one (see Zhu, 1956), since the latter could lead to undesirable comparisons within the taxonomic classes of the modified objects (rather than comparisons between the

object sets presented in the experiment; see Section 2.3).[1] We picked two simple relative gradable adjectives and two numerals, yielding four different pre-nominal sequences: *dada-de* 'big-DE', *xiaoxiao-de* 'small-DE', *yi-ge* 'one-CL', and *liang-ge* 'two-CL'. Correspondingly, we created different sets of novel objects (some of which were adapted from Miao, 2023) differing either in size or cardinality. See Table 1 for the full list of critical noun phrase items along with example novel objects used in the experiment.

Table 1. Noun phrase stimuli and example visual stimuli

Noun phrase stimuli			Example visual stimuli
[Adjectives] *dada-de* 'big-DE', *xiaoxiao-de* 'small-DE' **[Numerals]** *yi-ge* 'one-CL', *liang-ge* 'two-CL'	+	**[Novel nouns]** *qiaotong, xilu, shangpan, yunbei, kangchao, duhui, boquan*	

Aside from the novel noun phrases and objects described above, we also included four nouns familiar to children in the experiment: *qiqiu* 'balloon', *piqiu* 'ball', *jimu* 'toy brick', and *dangao* 'cake'. These items appeared as bare nouns accompanied by images without size or cardinality contrasts in trials that served as practice to familiarize children with the task. All noun phrases were embedded in sentences that constituted the speech context for different trials (see below) and were produced by a female native Mandarin speaker in a child-directed manner.

2.3. Design and predictions

The design of the experiment was modeled after Ferguson et al. (2014) and Syrett et al. (2019), both of which tested young children's interpretation of novel nouns using IPLP. During the task, participants were invited to identify objects based on a speech context that contained some object description. The experiment consisted of practice trials, test trials, and control trials (see below).

All trials started with an introduction phase. Participants saw two side-by-side pictures and listened to audio stimuli that invited them to examine the objects

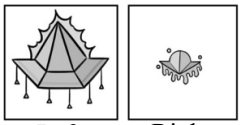
[1] As discussed in Zhu (1956), noun phrases modified by simple adjectives (e.g., *bai(-de) zhi* 'white paper') often identify subcategories of the nouns, thereby facilitating contrasts with other subcategories (e.g., *hong(-de) zhi* "red paper"), whereas this effect tends to be absent in noun phrases with complex (e.g., reduplicated) adjectives.

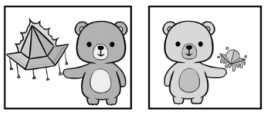
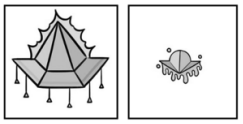
on the pictures. Importantly, in this phase children also heard the naming of the object(s) on one of the pictures (e.g., *Yi-bian shi qiaotong. Ling yi-bian bushi!* 'One side is qiaotong, and the other is not!' where *qiaotong* is the novel word). When novel words were used for naming, they were compatible with the object(s) on either picture. This was true even when the two pictures differed in cardinality, since nouns in Mandarin are not typically marked for plurality.

The dialogue phase followed, during which children heard some description about the object(s) named in the previous phase. The description here was embedded in a narrative involving an animal taking away the named object(s), to make it explicit that the description was about the exact same object(s), rather than some new object(s) sharing the same name. Correspondingly, participants first saw a blank screen and then two animals holding the object(s) from the previous phase in their hands. When the animal images were shown, children heard an utterance including either adjective- or numeral-containing novel noun phrases (ADJ and NUM conditions henceforth) twice in test and control trials (e.g., *Ta na zhe dada-de/yi-ge qiaotong* 'It's holding a big / one qiaotong!'). In practice trials, they heard utterances with bare nouns.

The final test phase began after the dialogue phase, with the side-by-side pictures from the introduction phase presented again. In practice and test trials, children were asked to identify the object(s) described in previous phases via a *where*-question that contained a bare noun (e.g., *Qiaotong zai naer?* 'Where is qiaotong?'). Table 2 below presents an example test trial. By contrast, in control trials, children were invited to find the object(s) with a different label than the one used in the earlier description—that is, they were asked to find the undescribed object(s). For instance, if the description was about *qiaotong*, children would be asked *Boquan zai naer?* 'Where is boquan?'. Such control trials were necessary since they helped us determine whether children integrated nouns with the pre-nominal elements in test trials—for instance, if they simply ignored the nouns, their performance should be identical in test and control trials. All *where*-questions were repeated twice, followed by a 3-second silence window.

Table 2. An example test trial from the ADJ condition

Phase	Duration	Audio stimuli	Visual stimuli
Introduction	9s	*Wa! Ni kan dao le ma? Yi-bian shi qiaotong. Ling yi-bian bushi!* 'Wow! Can you see it? One side is qiaotong, and the other is not!'	 Left Right

Dialogue	8s	*Ei? Bu jian le! You-zhi xiaoxiong na zou le qiaotong!* 'Oh? It's gone! A little bear took the qiaotong away!'	(Blankness)
	7s	*Ni kan! Ta na zhe dada-de qiaotong! Ta na zhe dada-de qiaotong!* 'Look! It's holding a big qiaotong! It's holding a big qiaotong!'	 Left Right
	0.5s	Silence	(Blankness)
Test	8s	*Hai you ne! Kan! Qiaotong zai naer? Qiaotong zai naer ne?* 'There is more! Look! Where is qiaotong? Where is qiaotong?'	 Left Right
	3s	Silence	

The experiment adopted a within-subject design. All participants were tested on the same types of trials and on both conditions (ADJ and NUM). Each participant first saw a practice trial to get familiar with the task, followed by two test trials (one for each condition) and one control trial from one condition. This sequence was then repeated, with a control trial from the other condition. Across trials, we counterbalanced the following factors: (i) the position (left/right) of the more salient (bigger size or greater cardinality) image; (ii) the position of the object(s) inquired in the test phase; and (iii) the pre-nominal item used in each test or control trial. In total, each child completed eight trials, typically within six minutes.

This design allowed us to examine children's integration between the pre-nominal elements and the nouns presented in the dialogue phase of test trials. Since the nouns were bare in the test phase, the information guiding their selection must come from the dialogue phase. If children successfully constructed and

integrated the representations of those pre-nominal elements and nouns in the dialogue phase, we predicted that they would (i) look more to the described object(s) in the test phase of test trials; and crucially, (ii) show a distinct looking pattern in the test phase of control trials. To reiterate, we should expect consistent failure if the integration process *per se* was difficult for children at our test age. Otherwise, difficulties should be observed selectively, only in items whose representations are challenging to build. Specifically, given the current design, the ADJ condition was predicted to be more difficult in general, as building the semantic representations of the two adjectives required establishing a proper comparison class (i.e., making a comparison between the two presented objects, rather than within their respective taxonomic classes) that was not explicitly specified in the linguistic stimuli. Furthermore, even within each condition, we may expect item differences (e.g., reflecting variation in children's knowledge of particular items) if children's integration ability was not uniformly vulnerable.

2.4. Procedure

Children were tested one at a time in a sound-proof room together with their guardians. Guardians wore headphones playing masking music and were asked not to influence children during the experiment. All audio and visual stimuli were played via a TV in front of the children. After children were seated properly and started looking at the screen, the experimenter started the first trial. When a trial ended, a short animation with music would be played to maintain children's interest. The next trial would be initiated only if children fixated at the screen. Throughout the experiment, children's looking patterns were videotaped by a hidden camera.

2.5. Results

2.5.1. Data preparation and analysis

The recording files were manually coded frame-by-frame. Each file was segmented into frames at a rate of 50 frames per second. A trained researcher labeled the participant's looking behavior in each frame as "left", "right", or "away". 12.5% of the data were coded by a second coder, and the inter-coder reliability was over 90%.

For statistical analysis, we specifically extracted the test phase data from the first onset of novel noun to the end of each test or control trial, and converted the "left" and "right" labels to "described-object(s)" or "undescribed-object(s)", depending on the setup of each trial.[2] Following the practice adopted by Syrett et al. (2019), we excluded trials (44 out of 144) in which the proportion of "away"-frames exceeded 25% for analysis.

We conducted two sets of analyses. First, we analyzed children's overall

[2] Aside from the main analyses reported here, we also coded and analyzed the introduction phase data to rule out any a priori natural preferences for the described object(s).

looking preference via different logistic mixed-effects models, using the *lme4* R package (Bates et al., 2015). Second, we conducted cluster-based permutation analyses (Dink & Ferguson, 2015) to evaluate children's looking-while-listening patterns, utilizing the *eyetrackingR* package (Dink & Ferguson, 2015). Analyses were primarily focused on test trial data. Specifically, we examined whether children looked more to the described object(s) than the undescribed one(s) in test trials, and also whether the patterns differed between test and control trials.

Since both ADJ and NUM conditions tested two distinct items, we first conducted two item analyses. We fit the test trial data in the two conditions to separate logistic mixed-effects models, predicting described-object(s) looking (coded as 1, and 0 otherwise) from individual item (sum-coded). For both models (and other models in this study), we included random effects that did not lead to convergence issues. If results indicated a difference between items, we divided the condition into two sub-conditions in the following analyses.

For overall looking preference analyses, we submitted the test trial data in each condition to different intercept-only models, with the dependent variable being children's looking to the described-object(s) (1/0). Significant intercepts would suggest robust overall looking preferences. As for the comparisons to the control trials, the models included an additional predictor, namely trial type (i.e., test vs. control, also sum-coded). Significant main effects of the predictor would indicate toddlers' distinct looking preferences between the two types of trials.

For looking-while-listening pattern analyses, we first averaged the adjacent two data points into 40-ms bins (a bin size comparable to De Carvalho et al., 2021). To determine whether there were time windows during which children exhibited reliably looking differences between the described and undescribed objects in test trials, we compared children's proportions of looks to the described object(s) to an artificial dataset where these proportions were set to 0.5 throughout (see Steil et al., 2021). Based on independent t-tests run on each time bin as well as 1,000 simulations, the analyses would return time windows where significant effects were detected. Comparisons between test and control trials were directly conducted on their data.

2.5.2. Experimental results

For item analyses, the model for the ADJ condition revealed no significant effects (all p's > 0.1), while the main effect of individual item approached significance ($p = 0.065$) in the NUM model. Since children's looking patterns in trials using different numerals showed a trend towards different patterns, we divided the NUM condition into NUM-ONE and NUM-TWO conditions in the subsequent analyses.

Now let's turn to the results of the overall looking preference analyses. Model results were consistent across different comparisons. In both ADJ and NUM-ONE conditions, children's looks to the described object(s) in test trials were not significantly different from their looks to the undescribed one(s), nor from the described one(s) in control trials (all p's > 0.1). Nevertheless, both comparisons

were significant in the NUM-TWO condition, indicated by the significant intercept in the corresponding intercept-only model and the significant main effect in the model including trial type as a predictor respectively (both p's < 0.05). The directions of both effects were consistent with the predictions of successful integration (see Section 2.3). In other words, successful integration was observed in the NUM-TWO condition, but not the others.

The results of different cluster-based analyses further corroborated the above patterns. There were only significant clusters in the NUM-TWO condition. The comparison between the test trial and chance-level (0.5) datasets revealed a window (from frame 196 to frame 270) in which children reliably looked more to the described object(s) (p = 0.012). The comparison between test and control trials also indicated children's distinct looking patterns from frame 218 to frame 276 (p = 0.008). Both significant windows started within 1.6 seconds after the first onset of the novel noun (frame 140) and extended to approximately the end of the first *where*-question in the test phase.

Figure 1 below plots children's overall proportions of described-object(s) looks across conditions, calculated by dividing the number of described-object(s) looks by the total number of both described- and undescribed-object(s) looks. Figure 2 visualizes children's looks over time via loess smoothed curves. Gray shaded boxes in the upper half (0.5-1) of the y-axis indicate significant differences between the test trial and chance-level (0.5) datasets, while those spanning the full y-axis suggest significant differences between test and control trials. Chance levels are indicated by dot-dash lines in both figures.

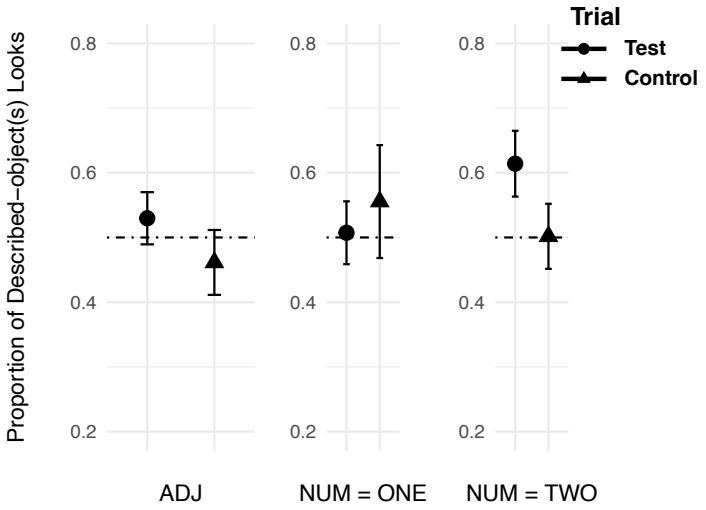


Figure 1. Participants' overall looks to the described object(s). Error bars show standard errors of by-subject means.

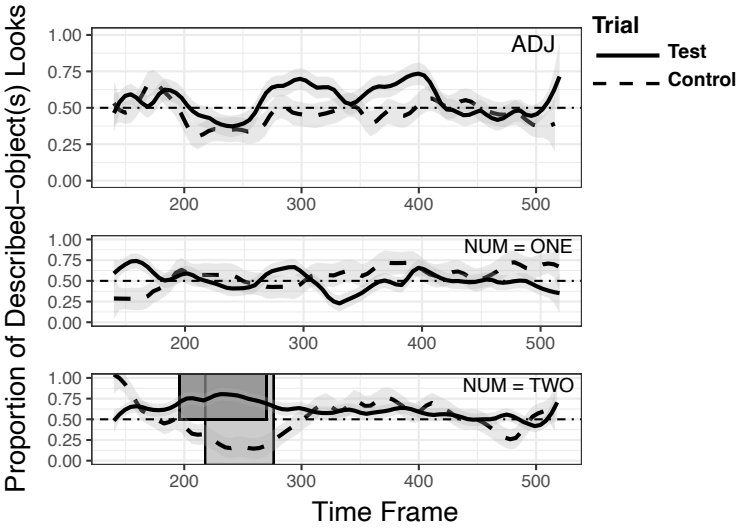


Figure 2. Participants' looks to the described object(s) over time. Light shading around the solid and dashed lines represents the 95% confidence intervals. Gray shaded boxes indicate time windows with significant effects.

3. Discussion

We tested Mandarin-learning 30-month-olds' ability to use different familiar adjectives and numerals to resolve novel noun reference in an IPLP experiment. Children were able to efficiently identify the correct referents through computing the integrated meanings of novel noun phrases containing the Mandarin numeral 'two', but failed to do so when the integration implicated relative gradable adjectives 'big/small' or the numeral 'one'.

First of all, our study shows that even for children as young as 30 months, the ability to efficiently integrate elements in noun phrases is present. Provided that they are able to build the representations of the words to be integrated (e.g., the numeral 'two' in our experiment), they can restrict the referent set properly even in an experimental setting that monitors their online processing under time pressure. Our study therefore goes beyond previous studies (e.g., Davies et al., 2021; Fernald et al., 2010; Ninio, 2004; Thorpe et al., 2006) in clearly demonstrating the ability of integration in children below three, via both their overall looking preferences and online reactions. This finding also suggests that the computation underlying linguistic composition may be a fundamental part of human cognition.

That being the case, the apparent integration failures in the ADJ and NUM-ONE conditions in our study (and possibly the ones in previous studies) may be better understood as children's difficulty in building the representations to be integrated. This is also in line with the selective difficulty reported in Redolfi &

Melloni (2025). Let's first consider our ADJ condition. As pointed out in Sections 1 and 2.3, building the semantic representations of our adjective elements requires establishing a comparison class. Since our linguistic stimuli did not make explicit the objects being compared, participants would need to properly restrict the entities under comparison (namely, the presented novel objects) based on the visual context, and also suppress comparisons within the objects' own taxonomic classes (i.e., interpreting the 'big' in 'big qiaotong' as larger than typical members of the qiaotong category). In light of this reasoning, such difficulty is predicted to disappear when the linguistic stimuli help restrict the object set being compared—children should perform better, as observed in our NUM(-TWO) condition, where the need to establish a comparison class never arises.

Now we turn to the asymmetry between the two numeral conditions, which may seem counter-intuitive at first sight. Children typically acquire 'one' before 'two', and for Mandarin specifically, the semantics of 'two' is reported to be not fully mastered at 30 months (see Le Corre et al., 2016). One possibility of children's failure in the NUM-ONE condition is that at this age, they allow the weak, indefinite reading of 'one', which is an option in Mandarin as well as many languages around the world (see Dryer, 2013). Unlike adults, they cannot efficiently utilize the context contrasting cardinality to access the appropriate non-weak reading. If this is the case, then for the children in our experiment, the critical noun phrases in the NUM-ONE condition were equivalent to *a/an N* (rather than *one N*). It has also been documented in the literature that children do accept the use of *a/an N* when the context has more than one object (e.g., Barner et al., 2009). That is to say, it is possible that children in our experiment thought both pictures were compatible with the NUM-ONE descriptions, and therefore failed the task. Regarding children's success with the numeral 'two' despite their incomplete semantic knowledge, we wish to highlight that the success of our NUM-TWO condition requires only the understanding that 'two' does not mean 'one', which would be sufficient to identify the object set with cardinality greater than one. This is also in line with the general finding about numeral acquisition in the literature—that is, children typically understand the semantic distinctions between different numerals before fully acquiring their exact meanings (see Condry & Spelke, 2008, among many others).

To conclude, our study reveals young children's ability to resolve novel noun reference by integrating elements in noun phrases. We also show that this ability is modulated by the difficulty in efficiently constructing proper representations of the words to be integrated. Seen in this light, children's performance—successful or not—provides evidence for their sensitivity to subtle linguistic distinctions.

References

Barner, David, Chow, Katherine, & Yang, Shu-Ju. (2009). Finding one's meaning: A test of the relation between quantifiers and integers in language development. *Cognitive Psychology*, *58*(2), 195–219.

Bates, Douglas, Mächler, Martin, Bolker, Ben, & Walker, Steve. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1).

Condry, Kirsten F., & Spelke, Elizabeth S. (2008). The development of language and abstract concepts: The case of natural number. *Journal of Experimental Psychology: General*, *137*(1), 22–38.

Davies, Catherine, Lingwood, Jamie, Ivanova, Bissera, & Arunachalam, Sudha. (2021). Three-year-olds' comprehension of contrastive and descriptive adjectives: Evidence for contrastive inference. *Cognition*, *212*, 104707.

De Carvalho, Alex, Dautriche, Isabelle, Fiévet, Anne-Caroline, & Christophe, Anne. (2021). Toddlers exploit referential and syntactic cues to flexibly adapt their interpretation of novel verb meanings. *Journal of Experimental Child Psychology, 203,* 105017.

Dink, Jacob W., & Ferguson, Brock. (2015). *EyetrackingR: An R library for eye-tracking data analysis* [Computer software].

Dryer, Matthew S. (2013). Indefinite articles. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Zenodo.

Ferguson, Brock, Graf, Eileen, & Waxman, Sandra R. (2014). Infants use known verbs to learn novel nouns: Evidence from 15- and 19-month-olds. *Cognition*, *131*(1), 139–146.

Fernald, Anne, Thorpe, Kirsten, & Marchman, Virginia A. (2010). Blue car, red car: Developing efficiency in online interpretation of adjective-noun phrases. *Cognitive Psychology*, *60*(3), 190–217.

Hirsh-Pasek, Kathryn, & Golinkoff, Roberta M. (1996). *The origins of grammar: Evidence from early language comprehension*. The MIT Press.

Kennedy, Christopher. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, *30*(1), 1–45.

Kennedy, Christopher. (2015). A "de-Fregean" semantics (and neo-Gricean pragmatics) for modified and unmodified numerals. *Semantics and Pragmatics*, *8*.

Le Corre, Mathieu, Li, Peggy, Huang, Becky H., Jia, Gisela, & Carey, Susan. (2016). Numerical morphology supports early number word learning: Evidence from a comparison of young Mandarin and English learners. *Cognitive Psychology*, *88*, 162–186.

Maris, Eric, & Oostenveld, Robert. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190.

Medin, Douglas L., & Shoben, Edward J. (1988). Context and structure in conceptual combination. *Cognitive Psychology*, *20*(2), 158–190.

Miao, Miao. (2023). *Early acquisition of classifiers in Mandarin Chinese: A preferential looking study* [Doctoral dissertation]. Tsinghua University.

Ninio, Anat. (2004). Young children's difficulty with adjectives modifying nouns. *Journal of Child Language*, *31*(2), 255–285.

Redolfi, Michela, & Melloni, Chiara. (2025). Processing adjectives in development: Evidence from eye-tracking. *Journal of Child Language*, *52*(2), 270–293.

Rothstein, Susan. (2013). A Fregean semantics for number words. In Maria Aloni, Michael Franke, & Floris Roelofsen (Eds.), *Proceedings of the 19th Amsterdam Colloquium* (pp. 179–186). Universiteit van Amsterdam.

Steil, Jessica N., Friedrich, Claudia K., & Schild, Ulrike. (2021). No evidence of robust noun-referent associations in German-learning 6- to 14-month-olds. *Frontiers in Psychology*, *12*, 718742.

Syrett, Kristen, LaTourrette, Alexander, Ferguson, Brock, & Waxman, Sandra R. (2019). Crying helps, but being sad doesn't: Infants constrain nominal reference online using known verbs, but not known adjectives. *Cognition*, *193*, 104033.

Thorpe, Kirsten, Baumgartner, Heidi, & Fernald, Anne. (2006). Children's developing ability to interpret adjective-noun combinations. In David Bamman, Tatiana Magnitskaia, & Colleen Zaller (Eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development* (Vol. 2, pp. 631–642). Cascadilla Press.

Zhu, Dexi. (1956). Xiandai Hanyu xingrongci yanjiu [Studies of adjectives in modern Chinese]. *Yuyan Yanjiu* [Studies in Language and Linguistics], (1), 83-111.