Cognition 176 (2018) 215–219

Contents lists available at ScienceDirect

Cognition

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

Brief article

Individual differences in nonverbal prediction and vocabulary size in infancy

Tracy Reuter[a,*], Lauren Emberson[a], Alexa Romberg[b], Casey Lew-Williams[a]

[a] Department of Psychology, Princeton University, USA
[b] Department of Human Development and Quantitative Methodology, University of Maryland-College Park, USA

ARTICLE INFO

Keywords:
Prediction
Language development
Individual differences
Nonverbal
Visual
Learning

ABSTRACT

Children who generate and update verbal predictions have larger vocabularies, suggesting that prediction may be a mechanism that supports language learning. We hypothesize that this relation is not confined to the domain of language, but instead signals a broader individual difference in information processing. To investigate this possibility, we tested infants ($n = 50$) in the early stages of vocabulary development (12–24 months) on their ability to generate and update nonverbal, visual predictions. In an eye-tracking task, a central fixation reliably preceded a peripheral target. Then, halfway through the experiment, the peripheral target began appearing on the opposite side. We assessed infants' proficiency in initiating anticipatory eye movements before and after the switch, and found that infants with larger vocabularies did not generate more predictions overall, but were more efficient in updating predictions to the new target side. These findings establish a link between nonverbal prediction and vocabulary in infancy, and suggest a promising means of addressing whether or not prediction abilities are causally related to language learning.

1. Introduction

Human processing of complex information is facilitated by prediction (Bar, 2007; Summerfield & de Lange, 2014). Humans make predictions in many domains, such as vision (Rao & Ballard, 1999; den Ouden, Friston, Daw, McIntosh, & Stephan, 2009; Summerfield & de Lange, 2014), locomotion (Wolpert, Miall, & Kawato, 1998; Wolpert, Ghahramani, & Flanagan, 2001), and language (Rabagliati, Gambi, & Pickering, 2016). In language, prediction enables efficient processing among both adults and children, allowing listeners to keep pace with the rapid information flow of speech (DeLong, Urbach, & Kutas, 2005; Kutas, DeLong, & Smith, 2011; Borovsky, Elman, & Fernald, 2012; Pickering & Garrod, 2013).

In addition to its role in language processing, prediction may also be a mechanism that facilitates language learning. In error-based models of language learning, learners compare predicted input with actual input to gain information about the structure of their language (Chang, Dell, & Bock, 2006; Elman, 1990; Pickering & Garrod, 2013). For example, a child might expect to hear the word 'mouses' but instead hear 'mice,' and update future predictions accordingly (Ramscar, Dye & McCauley, 2013). There are two types of evidence that these models may be valid descriptions of learning. First, it is well-established that children generate predictions during language processing. They are capable of drawing upon many types of linguistic information to anticipate what a speaker is likely to say next, such as phonology (Swingley, Pinto, & Fernald, 1999), semantics (Fernald, Zangl, Portillo, & Marchman, 2008; Fernald, Thorpe, & Marchman, 2010; Mani & Huettig, 2012), morphosyntax (Lew-Williams & Fernald, 2007; Borovsky et al., 2012; Lukyanenko & Fisher, 2016), and speakers' intentions (Kidd, White, & Aslin, 2011). Second, there are individual differences in the extent to which children generate verbal predictions, and these differences are related to children's language proficiency. Compared to children with smaller vocabularies, children with larger vocabularies are more likely to generate predictions in light of new linguistic information (Nation, Marshall & Altman, 2003; Borovsky et al., 2012; Mani & Huettig, 2012). Thus, in line with error-based models of learning, children who generate more verbal predictions and update those predictions efficiently have more advanced language abilities.

This research suggests that children can use multiple sources of information to anticipate downstream words and revise predictions as new linguistic information arrives. Although findings of this nature establish a link between prediction and language learning, they present an interpretational problem. There are a number of plausible explanations: One possibility is that verbal prediction is a capacity that supports vocabulary growth (see Elman, 1990). As reviewed above, prediction

* Corresponding author at: Department of Psychology, Princeton University, Princeton, NJ 08544, USA.
E-mail address: treuter@princeton.edu (T. Reuter).

https://doi.org/10.1016/j.cognition.2018.03.006
Received 20 March 2017; Received in revised form 24 January 2018; Accepted 6 March 2018
Available online 30 March 2018
0010-0277/

errors can be used to modify the learner's representations of their language, making future predictions increasingly accurate. In contrast, a second possibility is that verbal prediction is strictly an outcome, rather than a cause, of vocabulary growth (Rabagliati et al., 2016). That is, language users may only begin to generate predictions once they have a fair amount of linguistic knowledge. Evaluating these two possibilities, as well as intermediate views, will aid in understanding the role of prediction in language processing and learning.

To further examine the relation between prediction and language learning, we used two new approaches. First, we focused on infants between 1 and 2 years of age. Previous studies showing links between prediction and vocabulary have tested children between 2 and 7 years who already comprehend and produce multiword sentences. If prediction plays a role in supporting the initial stages of language learning, then infants' prediction abilities should already be linked to their budding linguistic knowledge. Second, in the current study we evaluated whether prediction as a domain-general capacity may be related to language learning. That is, we did not aim to replicate previously established relations between verbal prediction and vocabulary. Instead, based on views of prediction as a general capacity that is present in multiple domains and possibly interacts across domains (Bar, 2007; Lupyan & Clark, 2015), we examined relations between nonverbal (i.e., visual) prediction and vocabulary size. We reasoned that differences in nonverbal prediction, as compared to verbal prediction, are less likely to be the direct result of vocabulary differences. This cross-domain approach represents a new direction for understanding the relation between prediction abilities and language proficiency.

Our investigation of how infants make and update nonverbal predictions included two main hypotheses. First, we hypothesized that the *quantity* of predictions that infants generate in a nonverbal task would be linked to vocabulary. Among older children, those with larger vocabularies, as compared to those with smaller vocabularies, are more likely to make verbal predictions (Nation, Marshall, & Altmann, 2003; Borovsky et al., 2012; Mani & Huettig, 2012). We expected that this relation would hold earlier in development and apply to the domain of nonverbal prediction. Second, we hypothesized that the *quality* of infants' nonverbal predictions would be linked to vocabulary. In error-based models, learners update predictions when they encounter incongruent information (Chang et al., 2006). Assuming these models are relevant for explaining learning toward the beginning of life, we expected that infants with larger vocabularies would be more successful in updating nonverbal predictions after observing unexpected information.

To evaluate these hypotheses, we tested 12- to 24-month-old infants in a visual prediction task, using anticipatory eye movements (AEMs) as a measure of prediction. In a second eye-tracking task, we controlled for differences in infants' speed of visual processing. We compared performance on these tasks to infants' vocabulary size (MCDI). Together, we used these measures to evaluate whether and how nonverbal prediction abilities relate to infants' early language development.

2. Method

2.1. Participants

Participants were 50 infants (26 female) from monolingual English-speaking families who ranged in age from 12 to 24 months ($M = 18$, $SD = 3.5$). Infants were full-term and had no known vision or hearing impairments. We excluded an additional 14 infants from all analyses due to parental report of developmental delay (1), bilingual language exposure (2), fussiness such that less than 50% of trials were code-able (8),[1] or computer error (3). We excluded 6 of the 50 infants from visual processing task analyses due to computer error (4) or experimenter error (2). The Princeton University Institutional Review Board approved all research protocols, and a legal guardian provided informed consent for each infant.

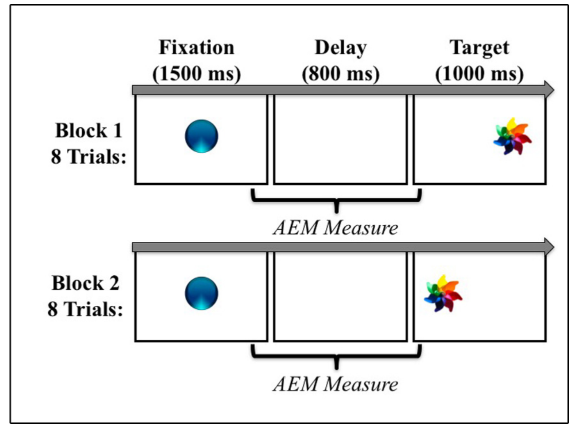

Fig. 1. Schematic of the prediction task. Infants saw two blocks of trials. In Block 1, the target always appeared on one side (e.g., right). In Block 2, the target always appeared on the opposite side (e.g., left). On each trial, we measured infants' anticipatory eye movements (AEMs), defined as looks to either target location during a time window from 200 ms before center fixation offset until 200 ms after target onset (AEM Measure).

2.2. Stimuli – Prediction task

On each trial, infants saw a central, looming fixation paired with a slide-whistle sound for 1500 ms. After an 800-ms delay, infants saw a peripheral, spinning target paired with another slide-whistle sound for 1000 ms (Fig. 1). Importantly, infants saw two blocks of trials. In the first block (trials 1–8), the target always appeared on one side, and in the second block (trials 9–16) its location switched sides. Block 1 target location was counterbalanced across infants.

On each trial, we measured infants' anticipatory eye movements (AEMs). As shown in Fig. 1, we conservatively defined AEMs as looks to either peripheral location during a time window from 200 ms before center fixation offset until 200 ms after target onset. This temporal window accounts for time needed to generate a saccade (Canfield, Smith, Brezsnyak, & Snow, 1997; Hallett, 1986; Matin, Shao, & Boff, 1993). AEMs were included in analyses regardless of infants' initial looking location.

Infants also saw a filler trial every 4 trials to maintain their attention. Fillers consisted of 5-s movies of a kaleidoscope paired with soft chimes. There was a 500-ms blank inter-trial interval.

2.3. Stimuli – Visual processing task

On each trial, infants saw a central fixation for 1000 ms, followed by a peripheral target for 1000 or 1250 ms (Fig. 2). Infants saw two types of trials. On gap trials, there was a 250-ms temporal gap between fixation offset and target onset. On overlap trials, there was a 250-ms temporal overlap between the fixation and the target. Unlike the prediction task, the target location did not follow a consistent pattern. Thus, infants were unable to accurately predict the target location. Trials appeared in one of two quasi-randomized orders, such that neither trial type (gap or overlap) nor target side (right or left) repeated for more than 3 trials sequentially. Fixation and target were stationary, and there were no auditory stimuli.

On each trial, we measured infants' reaction time (RT), defined as the time of the first target look occurring 200 ms or later after target onset (Fig. 2, "RT measure"). On overlap trials, the central stimulus

[1] We compared age and vocabulary measures for excluded and included infants, and found no differences in age [$t(8.58) = 0.06$, $p = 0.95$], MCDI comprehensive vocabulary size [$t(11.26) = -0.63$, $p = 0.54$], or MCDI productive vocabulary size [$t(11.71) = -0.40$, $p = 0.695$].

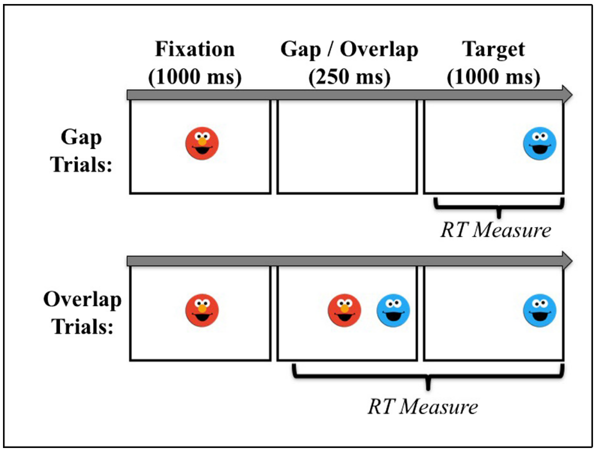

Fig. 2. Schematic of the visual processing task. Infants saw two types of trials. In gap trials, there was a 250-ms temporal gap between the fixation and the target. In overlap trials, there was a 250-ms temporal overlap between the fixation and the target. On each trial, we assessed infants' reaction time (RT), defined as the first target look during a time window from 200 ms after target onset (RT Measure).

competed for visual attention at target onset, so we expected infants' RT to be slower in overlap trials than in gap trials (Hood & Atkinson, 1993). We measured infants' visual processing speed as a difference score: overlap trials RT minus gap trials RT.

Infants also saw a filler trial every 8 trials to maintain their attention. Fillers consisted of 5-s movies of bubbles paired with child laughter. There was a 500-ms blank inter-trial interval. All stimuli are available in supplementary materials.

2.4. Procedure

Before testing, caregivers completed the MacArthur-Bates Communicative Developmental Inventory (MCDI). We administered the MCDI Words and Gestures for infants 12–15 months old, and MCDI Words and Sentences for infants 16–24 months old, and standardized infants' scores in accordance with the MCDI technical manual (Fenson et al., 2007). During testing, each infant completed the prediction task, and then immediately completed the visual processing speed task. Infants sat on their caregiver's lap, and caregivers wore occluded glasses to avoid influencing the infants' behavior. Infants were approximately 70 cm from a large, flat-screen television monitor (68 × 122 cm). The experimenter recorded infants' eye movements via a video camera below the screen. After testing, a research assistant coded videos of infants' eye movements, indicating the direction of looks as: left, center, right, away, or shifting. Another research assistant also coded 25% of the videos. Inter-coder reliability was 98%.

3. Results

We hypothesized that infants with larger vocabularies would make more predictions and would update predictions more readily than infants with smaller vocabularies. To evaluate these hypotheses, we correlated infants' productive vocabulary percentile (MCDI) and anticipatory eye movements (AEMs) during the prediction task. We excluded trial 1 from all analyses, as there was no basis for infants to generate predictions therein. We also compared high-vocabulary and low-vocabulary groups, based on a median split in MCDI scaled scores. Finally, we analyzed visual processing speed to evaluate alternative explanations for individual variability in infants' performance on the prediction task.

3.1. Prediction task

To evaluate whether infants with larger vocabularies generated more predictions overall (regardless of accuracy), we first correlated infants' vocabulary size with their proportion of total trials with an AEM to either location. There was striking variation in both vocabulary size (MCDI percentile range = 5–99, $M = 42$, $SD = 28$) and proportion of trials with an AEM (range = 0–100, $M = 55$, $SD = 29$). Including all infants, regardless of how many AEMs they generated, we found no correlation between infants' vocabulary size and overall proportion of trials with an AEM [$r(48) = -0.12$, $p = 0.39$]. Examining each Block separately, there was no correlation in Block 1 [$r(48) = -0.03$, $p = 0.82$] nor in Block 2 [$r(47) = -0.17$, $p = 0.23$]. Thus, contrary to our first hypothesis, infants with relatively larger vocabularies were not more likely to anticipate the target's appearance.

To test our second hypothesis, that infants with larger vocabularies would update predictions more readily, we classified AEMs in Block 2 as directed toward the *novel* location (i.e., the Block 2 target location) or the *familiar* location (i.e., the Block 1 target location). Trial 9 was excluded from this analysis, because this was the first trial in which the target appeared in the novel location. Thus, infants had no basis for directing an AEM to the novel target location on this trial. Infants who did not make any AEMs in Block 2 ($n = 8$) or who did not have any code-able trials ($n = 1$) were not included in this analysis. For each infant, we calculated the proportion of their AEMs that were to the novel location (i.e., novel AEMs divided by total AEMs).

We found a significant correlation between infants' vocabulary size and the proportion of their AEMs to the novel location [$r(39) = 0.34$, $p = 0.028$]. Consistent with our second hypothesis, when making an AEM in Block 2, infants with larger vocabularies were more likely to make an AEM to the novel target location than infants with smaller vocabularies (Fig. 3). By comparison, in Block 1, 98% of infants' AEMs were directed to the correct peripheral target location, and there was no correlation between correct AEMs in Block 1 and vocabulary size [$r(42) = 0.24$, $p = 0.12$].

To examine the time-course of prediction updating, we analyzed AEMs trial by trial in Block 2. We divided infants into high-vocabulary and low-vocabulary groups based on a median split in MCDI percentile scores (high vocabulary: $n = 22$, $M = 64$, $SD = 19$; low vocabulary: $n = 19$, $M = 17$, $SD = 9$). Then, on each trial we calculated the proportion of AEMs to the novel location as the number of infants who made an AEM to the novel location divided by the number of infants who made AEMS to either location (Fig. 4). A binomial test was used to test this proportion against a chance value of 0.5. The low-vocabulary group only performed above chance on trial 16 ($p = 0.039$), making significantly more AEMs to the novel than the familiar location. The high-vocabulary group was marginally above chance in trial 13

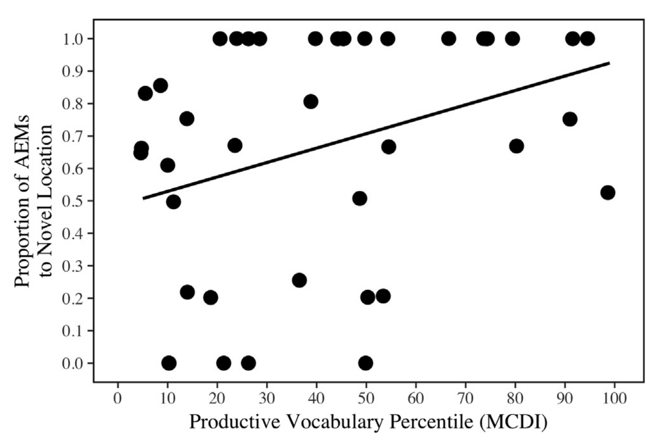

Fig. 3. Proportion of AEMs to the novel target location in Block 2. Line indicates linear regression and points represent individual subjects (n = 41).

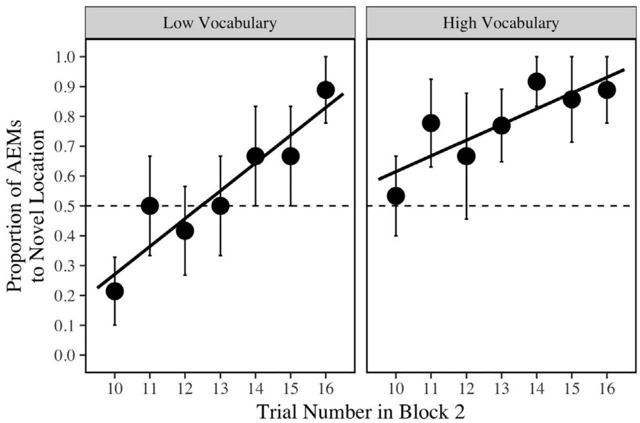

Fig. 4. Proportion of AEMs to the novel target location, by trial in Block 2, for high vocabulary infants (n = 22) and low vocabulary infants (n = 19). Lines indicate linear regression and bars represent standard error of the mean.

($p = 0.092$), above chance in trial 14 ($p = 0.006$), and above chance in trial 16 ($p = 0.039$).

To assess whether nonverbal prediction varied as a function of age, we repeated the above correlations with infants' age in months as a factor. We found no correlation between infants' age and overall proportion of trials with an AEM [$r(48) = -0.25$, $p = 0.08$]. There was no correlation in Block 1 [$r(48) = -0.23$, $p = 0.11$] or in Block 2 [$r(47) = -0.22$, $p = 0.13$]. Finally, there was no correlation between infants' age and the proportion of their AEMs to the novel location in Block 2 [$r(39) = 0.10$, $p = 0.52$].

3.2. Visual processing task

The gap-overlap task allowed us to examine whether individual differences in AEMs were due to differences in visual processing speed. First, we confirmed that infants were slower to shift their eyes to the target on overlap trials when there was a competing stimulus than on gap trials when there was no competing stimulus [gap trials: $M = 408$ ms, $SD = 31$; overlap trials: $M = 509$ ms, $SD = 74$; two-tailed $t(42) = -10.4$, $p < .001$]. Next, we computed RT difference scores for each infant (i.e., overlap trials minus gap trials) and used this difference score as the measure of visual processing speed independent of prediction. We found no correlation between infants' RT difference scores and their proportion of AEMs overall [$r(40) = -0.09$, $p = 0.56$] and no correlation between infants' RT difference scores and their proportion of AEMs to the novel target location in Block 2 of the prediction task [$r(33) = 0.18$, $p = 0.29$]. Together, these findings suggest that infants' performance on the prediction task was not related to variation in orienting to the dynamics of visual events.

4. Discussion

A growing body of research demonstrates a link between verbal prediction and language learning (Nation et al., 2003; Borovsky et al., 2012; Mani & Huettig, 2012), but the nature of this relation remains uncertain (Rabagliati et al., 2016). Specifically, is prediction a cause or a consequence of language learning? While the present study does not license causal inferences, our approach represents one step toward addressing this directionality puzzle by investigating nonverbal, visual prediction in infancy.

In particular, the present findings make two novel contributions to the current understanding of prediction and language learning. First, we demonstrate a link between infants' nonverbal prediction and vocabulary. Nonverbal prediction, as compared to verbal prediction, is less likely to be the direct result of language learning. Thus, our findings suggest that language learning may be supported in part by domain-general prediction abilities (see Bar, 2007; Summerfield & de Lange, 2014). Second, we show that the ability to update predictions in infancy – and to not perseverate with previously learned predictions – may be particularly important for learning. Infants with larger vocabularies did not generate more predictions overall, but they updated predictions quickly in light of new information. Together, these findings suggest that prediction capacities beyond the domain of language are relevant for understanding early vocabulary growth, and that the quality of predictions (i.e., updating) rather than quantity of predictions (i.e., overall tendency to predict) might support vocabulary growth.

How might infants' abilities to flexibly update predictions support language learning? Updating predictions to adapt to changes in the environment is a central feature of error-based learning models (e.g., Rescorla & Wagner, 1972). In such models, predictions are not static. Rather, predictions are continuously updated in light of new information, so that the learner can generate more accurate predictions in the future. Updating predictions may enable infants to better represent the shifting statistics of the input, incrementally generating more precise predictions as they gain more information. In contrast, maintaining the same predictions throughout learning may block infants from processing and encoding novel information. Individual differences in infants' abilities to update predictions may operate from early in infancy and contribute to gains in a variety of domains, including language.

However, it is not clear what causes early differences in prediction abilities. One possibility is that learning from more variable input encourages infants to update predictions more frequently. Dynamic, rich experience with objects, people, and language could reinforce the benefits of attending to and learning from new information. Supporting this view in a laboratory context, Romberg and Saffran (2013) found that infants who received variable input in an audio-visual learning task updated predictions more readily than those who received deterministic input. Similarly, Kovács and Mehler (2009) found that bilingual infants, as compared to monolinguals, rapidly updated predictions in a nonverbal task. In both experiments, there was variation in prediction updating, much like the present findings; some infants never updated predictions despite the fact that their predictions were consistently incorrect. Thus, learning from more variable input may encourage prediction updating. Another possibility is that the relation between prediction and language learning is mediated by related cognitive capacities, such as executive function (Kovács & Mehler, 2009) or working memory (Mani & Huettig, 2013; Huettig & Janse, 2016). Further experiments are needed to fully characterize the connections between prediction and language learning, by taking into account both linguistic and nonlinguistic factors and by considering plausible third variables. Here, we suggest that a more general framework of predictive information processing, especially one that captures how learners update predictions, will be fruitful in understanding how infants begin learning patterns in the world, including language.

Acknowledgements

This research was funded by grants from the National Institute of Child Health and Human Development to Casey Lew-Williams (R03HD079779) and to Lauren Emberson (R00HD076166-02). We thank the participating families and members of the Princeton Baby Lab. We are especially grateful to Jean Bellamy and Fernanda Fernandez for coding.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2018.03.006.

Stimuli, data, and R code are available on Dataverse (doi: http://dx.doi.org/10.7910/DVN/GFVHHS).

References

Bar, M. (2007). The pro-active brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences, 11*, 280–289.

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology, 112*(4), 417–436.

Canfield, R. L., Smith, E. G., Brezsnyak, M. P., & Snow, K. L. (1997). Information processing through the first year of life: A longitudinal study using the visual expectation paradigm. *Monographs of the Society for Research in Child Development, 62*, 1–45.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*(2), 234.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117–1121.

den Ouden, H. E. M., Friston, K. J., Daw, N. D., McIntosh, A. R., & Stephan, K. E. (2009). A dual role for prediction error in associative learning. *Cerebral Cortex, 19*(5), 1175–1185.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14*(2), 179–211.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *Macarthur-bates communicative development inventories: User's guide and technical manual* (2nd ed.). Baltimore, MD, USA: Brookes.

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

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernandez, & H. Clahsen (Eds.). *Developmental psycholinguistics: Online methods in children's language processing*. Amsterdam, Netherlands: John Benjamins Publishing Company.

Hallett, P. E. (1986). Eye movements. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.). *Handbook of perception and human performance*. New York, NY, USA: Wiley.

Hood, B. M., & Atkinson, J. (1993). Disengaging visual attention in the infant and adult. *Infant Behavior and Development, 16*(4), 405–422.

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience, 31*(1), 80–93.

Kidd, C., White, K. S., & Aslin, R. N. (2011). Toddlers use speech disfluencies to predict speakers' referential intentions. *Developmental Science, 14*(4), 925–934.

Kovács, A. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences, 106*(16), 6556–6560.

Kutas, M., DeLong, K. A., & Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In M. Bar (Ed.). *Predictions in the brain: Using our past to generate a future*. Oxford: Oxford University Press.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 18*(3), 193–198.

Lukyanenko, C., & Fisher, C. (2016). Where are the cookies? Two- and three-year-olds use number-marked verbs to anticipate upcoming nouns. *Cognition, 146*, 349–370.

Lupyan, G., & Clark, A. (2015). Words and the world: Predictive coding and the language-perception-cognition interface. *Current Directions in Psychological Science, 24*(4), 279–284.

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38*, 843–847.

Mani, N., & Huettig, F. (2013). Towards a complete multiple-mechanism account of predictive language processing. *The Behavioral and Brain Sciences, 36*(4), 365–366.

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics, 53*, 372–380.

Nation, K., Marshall, C. M., & Altmann, G. (2003). Investigating individual differences in children's real- time sentence comprehension using language-mediated eye movements. *Journal of Experimental Child Psychology, 86*, 314–329.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *The Behavioral and Brain Sciences, 36*(4), 329–347.

Rabagliati, H., Gambi, C., & Pickering, M. J. (2016). Learning to predict or predicting to learn? *Language, Cognition and Neuroscience, 31*, 94–105.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience, 2*(1), 79–87.

Ramscar, M., Dye, M., & McCauley, S. M. (2013). Error and expectation in language learning: The curious absence of mouses in adult speech. *Language, 89*(4), 760–793.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black, & W. F. Prokasy (Eds.). *Classical conditioning II: Current research and theory*. Appleton Century Crofts: New York, NY, USA.

Romberg, A. R., & Saffran, J. R. (2013). Expectancy learning from probabilistic input by infants. Frontiers in Psychology, *3*(JAN), 1–16.

Summerfield, C., & de Lange, F. P. (2014). Expectation in perceptual decision making: neural and computational mechanisms. Nature Reviews Neuroscience, (October), 1–12.

Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71*, 73–108.

Wolpert, D. M., Ghahramani, Z., & Flanagan, J. R. (2001). Perspectives and problems in motor learning. *Trends in Cognitive Sciences, 5*(11), 487–494.

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences, 2*(9), 338–347.