Young children predict the form of upcoming words during natural storytelling contexts.

Anthony Yacovone[1,2], Briony Waite[2], & Jesse Snedeker[2] – Tufts University[1], Harvard University[2]

Introduction. Adult language comprehension is predictive.[1] When we listen to a story, we do not merely take in each new sound as it is produced and map it to possible word forms and meanings. We also anticipate (or *predict*) what might come next given our knowledge of our own language and of the unfolding story. Decades of psycholinguistic research have demonstrated predictive processing of this kind in adults. For example, studies using eye-tracking and EEG have shown that predictable words are read faster and evoke smaller neural responses (i.e. N400s) than less predictable words.[2,3] We also know that prediction can occur at both the levels of meaning and form, as evidenced by EEG studies that find reduced N400s to unexpected words when they overlap in meaning or word form with the expected target.[4–6] In short, there is ample evidence that prediction is a robust and prevalent feature of the mature language comprehension system.

We know very little, however, about the development of predictive language comprehension. The received wisdom in developmental psycholinguistics is that children often fail to use contextual information to constrain their moment-to-moment interpretation.[7–9] Thus, we might also expect that children would fail to predict during comprehension, as it requires the effective use of prior contextual information to anticipate what might come next. Indeed, the evidence for prediction in children is quite mixed: Most studies find that children can anticipatorily look at *sweet*, *edible* objects in a visual display after hearing “Mary baked the….” But there has been no definitive evidence for prediction of form in young children.[10–12] The present study begins to explore the question of when (and how) prediction develops by looking for evidence of form-based prediction in young children using an EEG task known as the *Storytime* paradigm. In this task, we record participants’ EEG while they simply listen to a coherent, naturally produced story with experimental manipulations spliced into it.

Design. After pre-registered exclusions, our sample included 28 English-speaking children (*Mean Age* = 5;8, *Range* = 5;0–6;9). For our story stimulus, we created a 30-minute cartoon of a children’s book and grafted a 2 × 3 factorial design into it (Table 1). Specifically, we altered a set of predictable and unpredictable target words such that children either heard the original word from the story (*cake*) or a non-word with a similar (*ceke*) or less-similar phonological form (*vake*). Predictability was assessed via cloze probabilities obtained from a different set of 45 children in a spoken cloze task using the same story (Predictable words: *Mean Cloze* = 57.9%, Unpredictable words: *Mean* = 4.5%). Note, the critical word was spliced in all conditions.

Predictions. In prior work with adults, evidence for form-based prediction manifested as smaller N400s for form-similar relative to less-similar violations—but only when adults could predict the original word.[6] Adults also produced late emerging, posterior positivities (P600s) to these form-based violations. P600s have been associated with the reprocessing of errors encountered during deep comprehension, and arguably reflect a core aspect of adults’ reasoning about erroneous language.[13] If children engage in form-based prediction, we expect to see reduced N400 responses to form-similar violations relative to less-similar violations in predictable but not unpredictable contexts. These predictions were tested using a pre-registered linear mixed effects model and pairwise testing of estimated marginal means for N400 amplitudes (μV, 500-800ms).

Results and Discussion. In predictable contexts, we found N400 effects for both violations (*cake* vs. *ceke*: $b = 2.04$, $t = 2.44$, $p = .014$; *cake* vs. *vake*: $b = 3.78$, $t = 4.55$, $p < .001$). Critically, the N400 response for the form-similar violation was reduced relative to the less-similar violation for predictable words (*ceke* vs. *vake*: $b = 1.74$, $t = 2.09$, $p = .037$). In unpredictable contexts, there were no differences between conditions, as even the baseline words evoked relatively large N400s (see Figure 1, Unpredictable waveforms). Interestingly, unlike in adults, we did not see P600s to these violations in children, generating new questions about error monitoring and how it changes across the lifespan. In sum, these findings provide the first clear evidence that young children *can* effectively use contextual information to predict the form of upcoming words.

References: **[1]** Kuperberg & Jaeger, 2016; **[2]** Staub, 2015; **[3]** Kutas & Hillyard, 1984; **[4]** Federmeier & Kutas, 1999; **[5]** DeLong et al., 2005; **[6]** Ito et al., 2016; **[7]** Trueswell et al. 1999; **[8]** Snedeker & Trueswell, 2004; **[9]** Yacovone et al., 2021; **[10]** Mani & Huettig, 2012; **[11]** Gambi et al., 2018; **[12]** Borovsky et al., 2012; **[13]** Kuperberg et al., 2020.

Table 1. Example sentences from the story.

Predictability	Word Type	Example Sentence
High Cloze	**Baseline word**	These hairs prove you were at the scene of the **crime** [kraɪm].
High Cloze	**Similar non-word**	These hairs prove you were at the scene of the **crame** [kreɪm].
High Cloze	**Less Similar non-word**	These hairs prove you were at the scene of the **nime** [naɪm].
Low Cloze	**Baseline word**	They are only found in a certain **river** [rɪvər] in Texas.
Low Cloze	**Similar non-word**	They are only found in a certain **ruver** [rʌvər] in Texas.
Low Cloze	**Less Similar non-word**	They are only found in a certain **piver** [pɪvər] in Texas.

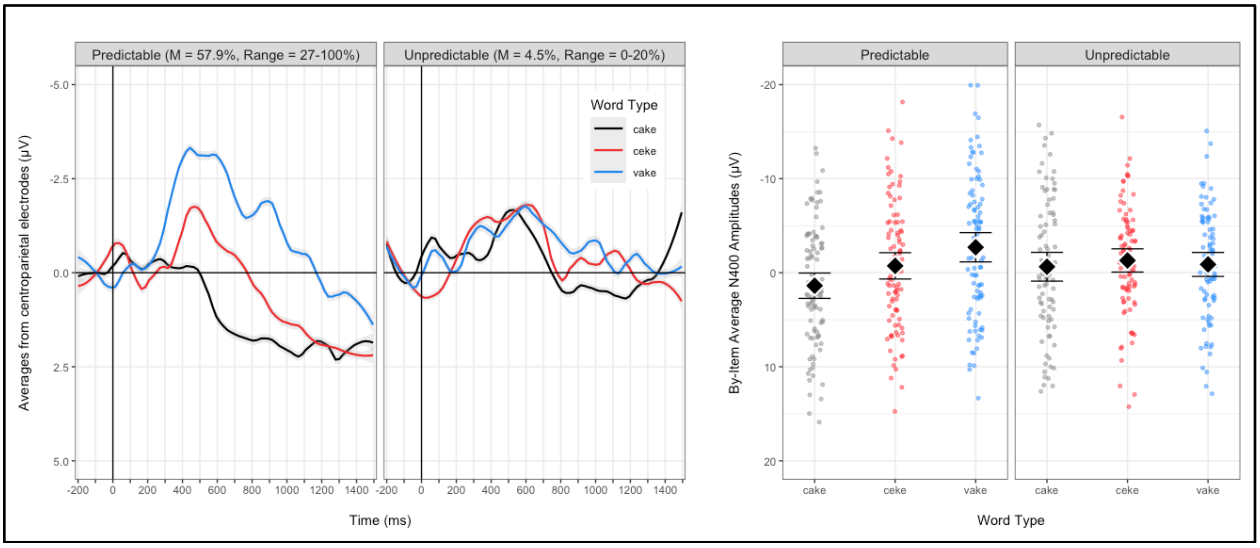

Figure 1. *Average waveforms and N400 amplitudes across predictability groups and word type.* In left panel, grand average waveforms from centroparietal electrodes (Cz, C3/4, CP1/2, Pz, P3/4) are plotted for predictable (left) and unpredictable (right) conditions. These waveforms were produced in R and then smoothed using local regression (loess) smoothing techniques. In the right panel, by-item mean N400 amplitudes for baseline words (cake, **black**), similar non-words (ceke, **red**), and less similar non-words (vake, **blue**) are plotted. Black diamonds represent the mean amplitude for the group.