To quantify each participant’s probability of guessing, we computed the proportion of signer-to-target (correct) and signer-to-distracter (incorrect) shifts for each child. Previous work using the Looking-while-Listening paradigm could not easily compute these values, since the task did not include a center fixation point. We then used a latent mixture model in which we assumed that the observed data (children’s initial shifts away from the signer) were generated by two processes (guessing and knowledge) that had different overall probabilities of success, with the “guessing group” having a probability of 50% and the “knowledge” group having a probability > 50%. The group membership of each participant was a latent variable inferred based on that participant’s proportion of correct signer-to-target shifts relative to the overall proportion of correct shifts across all participants (see Lee & Wagenmakers [2013] for a detailed discussion of this modeling approach). We then used each participant’s inferred group membership to weight participants *proportional* to our belief that they were guessing[[1]](#footnote-1). It is important to point out that we use this approach only in the analysis of RT, given our assumption that “guessing behavior” is integral to our measure of children’s mean accuracy in the VLP task, but not to our measure of mean RT which is based on correct trials.

In all of the Bayesian linear models[[2]](#footnote-2), we assume that each outcome variable (mean accuracies and RTs for each participant) is drawn from a Gaussian distribution with a mean, μ, and a standard deviation, σ. The mean is generated by a linear function consisting of an intercept term, α, which encodes the expected value of the outcome variable when the predictor is zero, and a slope term, β, which encodes the expected change in the outcome with each unit change in the predictor (i.e., the strength of association). We use vague priors for the intercept and the standard deviation, allowing the model to consider a wide range of plausible values.

We chose to use informative priors for the slope parameters in each model. Specifically, we used a truncated Gaussian distribution with a mean of zero and a standard deviation of one. Centering the distribution at zero is conservative and places the highest prior probability on a null association. By truncating the prior, we encoded our directional hypotheses for the associations between processing skills and age/vocabulary (i.e., that we predict that these relations should be null or improve with increasing age and larger vocabulary size). Finally, to constrain the range of plausible slope values in our model, we used previous research on the development of real-time language comprehension in children learning spoken language showing that the average gain for one month of development between 18-24 months in accuracy is ~0.016 and for RT is ~50 milliseconds (Fernald et al., 2008).

It is important to point out that our use of an informative prior does not affect parameter estimation. That is, if we substitute vague prior distributions, estimates of the strength of the associations between age/vocab and accuracy/RT are unchanged since there are enough data to overwhelm the uninformative priors. However, the use of an uninformative prior does directly influence Bayesian model comparison, which we use to quantify the strength of evidence for our linear models (i.e., the Bayes Factor[[3]](#footnote-3)). Intuitively, the use of an uninformative prior allows our models to predict any slope parameter, creating a situation where only a small amount of the prior probability is placed on a null relationship (i.e., where the slope is zero). Thus, the Bayes Factor, which is computed by taking the ratio of the prior and the posterior density when the slope is zero, is likely to show a preference for the null model, even when the data appear inconsistent with it (see Lee & Wagenmakers [2013]). The use of informative priors is an area of active debate in Bayesian statistical methods, but it has become a focus of recent work in Bayesian cognitive modeling (Lee & Vanpaemel, submitted).

1. Four children (ages: 18, 20, 22, and 25 months) had high posterior probability mass on guessing: posterior probabilities of 0.89, 0.86, 0.82, and 0.77, respectively (mean proportion signer-to-target scores for these participants were: 0.55, 0.46, 0.38, 0.33). [↑](#footnote-ref-1)
2. Models with categorical predictors were implemented in STAN (Stan Development Team, 2016). Models with continuous predictors were implemented in JAGS (Plummer, 2003). [↑](#footnote-ref-2)
3. The Bayes Factor can be interpreted as a measure of the relative strength of evidence one model (M1) over another model (M2): e.g., a of 5 means that the data are 5 times more likely given M1. [↑](#footnote-ref-3)