HHS Public Access
Author manuscript
Dev Sci. Author manuscript.

Using Language Input and Lexical Processing to Predict Vocabulary Size

Tristan Mahr and
University of Wisconsin–Madison, Department of Communication Sciences and Disorders, 1500 Highland Ave, Madison, WI 53705

Jan Edwards
University of Maryland

Tristan Mahr: tristan.mahr@wisc.edu

Abstract

Children learn words by listening to caregivers, and the quantity and quality of early language input predict later language development. Recent research suggests that word recognition efficiency may influence the relationship between input and vocabulary growth. We asked whether language input and lexical processing at 28–39 months predicted vocabulary size one year later in 109 preschoolers. Input was measured using adult word counts from LENA recordings. We used the visual world paradigm and measured lexical processing as the rate of change in proportion of looks to target. Regression analysis showed that lexical processing did not constrain the effect of input on vocabulary size. We also found that input and processing were more reliable predictors of receptive than expressive vocabulary growth.

Introduction

Vocabulary size is a robust predictor of language development. By amassing a large vocabulary of words, children learn morphosyntactic regularities (Marchman & Bates, 1994) and develop the phonological representations that subserve future reading skills (Walley, Metsala, & Garlock, 2003). Early delays in word learning can predict subtle group differences in vocabulary, syntax, and verbal memory assessment scores in the school-age years (Rescorla, 2009), and preschool expressive vocabulary and sentence complexity predict literacy development (Scarborough, 2009). Studying individual differences in word learning and how these differences emerge can inform our understanding of how differences in language ability emerge as well.

What does it take to learn words? In an abstract sense, children are word-learning machines, converting language input from caregivers into linguistic knowledge and spoken words of their own. In this formulation of the problem, children learn words by discovering and extracting regularities from their environment. Consequently, we may consider two axes that

MR. TRISTAN MAHR (Orcid ID : 0000-0002-8890-5116)

constrain word learning: the teaching environment and the child's ability to process and extract information.

There is a well-established literature documenting how environmental factors influence language development (Hoff, 2006). In particular, the amount of early language input from caregivers predicts language development later on (e.g., Hart & Risley, 1995; Hoff, 2003; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Rowe, 2012). In a landmark study, Hart and Risley (1995) found that children from different socioeconomic backgrounds experienced very different language environments. Children from higher SES families, compared to children from lower SES families, heard more words (tokens) and more different words (types) and showed better language outcomes. The reported "30 million word gap" in cumulative language input between high-SES and low-SES families has become a public health issue and the target for high-profile interventions (Shankar, 2014).

Although greater language input predicts greater language outcomes, there is no reason to think all words are equally informative. Research on input "quality" has examined whether some features of language input are more important and more informative than others for shaping language outcomes. Hoff (2003) found that features of caregivers' speech explained word-learning differences in two-year-olds from families with high school-educated mothers versus college-educated mothers. In particular, SES category explained little variance in vocabulary growth over and above caregiver mean-length of utterance, baseline vocabulary size, and child birth order. Hoff (2003) concluded that SES influences caregiver speech, but it is caregiver speech that drives language learning "by providing data to the child's word-learning mechanisms" (p. 1374). In this case, data from high-SES families included more word types and tokens and longer utterances.

A child's environment predicts language outcomes. Large changes in environment, measured by socioeconomic differences, predict group differences in the quantity and make-up of child-directed speech. But even within SES groups, families vary considerably in measures of child-directed speech (Weisleder & Fernald, 2013). There is also a quality-versus-quantity dimension to language input—not all words or utterances are equally educational—but language quantity may provide a reasonable approximation of the amount of high-quality learning experiences available to a child.

We now consider the other axis for word learning: a child's ability to process information. Speech perception, segmentation, and recognition skills measured in the first year predict vocabulary measures in the second and third years (see Cristia, Seidl, Junge, Soderstrom, & Hagoort, 2014 for a systematic review). For example, in a conditioned head-turn speech perception experiment with 6-month-olds, Tsao, Liu, and Kuhl (2004) found the number of trials needed to reach a criterion performance correlated with expressive vocabulary at 24 months ($r = -.48$). In this case, children who needed fewer trials to learn a contrast developed larger vocabularies 18 months later. Kuhl et al. (2008) found that event-related potentials at 7.5 months predicted vocabulary growth such that children who were more sensitive to native phonetic contrasts had larger vocabularies at 18 and 24 months compared to children who were more sensitive to non-native contrasts. Finally, a meta-analysis of six speech segmentation studies by Junge and Cutler (2014) found a median correlation of .33, 95% CI [.17, .48] between word segmentation measures and later language outcomes.

Infants who are better at extracting the sounds and shapes of words in their ambient language have larger vocabularies in toddlerhood. The ability to process words (lexical processing or word recognition) also predicts future language outcomes. This kind of processing is commonly assessed by eyetracking tasks which measure how quickly and reliably a child fixates on an image after hearing its label. In this paradigm, short-term and long-term effects have been found between processing and later outcomes. Fernald and Marchman (2012) found that late-talking toddlers with faster lexical processing at 18 months were more likely to move into the normal range of vocabulary scores by 24 months. Marchman and Fernald (2008) found that accuracy and speed of lexical processing at age two predicted language and working memory scores at age eight. Lany (2017) found a relationship between speed of lexical processing and novel word learning in 18-month-olds and 30-month-olds. Children who were faster at recognizing familiar words were also more accurate at recognizing novel words in a word-learning task.

Both language environment and ability to process speech can shape language outcomes, but only a few longitudinal studies have considered how these two factors work together. Newman, Rowe, and Bernstein Ratner (2015) examined how language input and processing at 7 months predicted vocabulary at 24 months. They found that amount of time listening to novel (unfamiliarized) words predicted vocabulary size, as did type-token-ratio of caregiver speech such that more repetitive speech predicted larger vocabularies. These two predictors jointly predicted vocabulary size, but did not significantly interact and were weakly correlated. Therefore, they concluded the learning environment and the child’s processing ability supported language development independently.

In eyetracking studies with older children, however, mediating relationships have been found between input and lexical processing. Hurtado, Marchman, and Fernald (2008) found that maternal talk at 18 months predicted lexical processing speed and vocabulary size at 24 months in 27 Spanish-learning children from predominantly low-SES families. Processing speed mediated the effect of input on vocabulary size, suggesting that maternal input provides practice for processing words and children who are more efficient at recognizing words learned more words. Vocabulary size, however, also mediated the effect of maternal speech on processing speed. Because processing and vocabulary size were measured at the same time, it is not clear which mediation path (or both, in some reciprocal effect) better explains the data. In short, these the measures are interrelated, with maternal input predicting future processing and vocabulary outcomes.

In a very similar study, Weisleder and Fernald (2013) studied 29 Spanish-learning children from low-SES families and found that processing and language input at 19 months of age predicted vocabulary size at 24 months. Lexical processing, however, mediated the effect of input on future vocabulary. The authors concluded that increased input affords more opportunities to practice recognizing words and that greater processing efficiency facilitated word learning. Another important question, however, is whether the beneficial effects of language input on later vocabulary size are *constrained* by the child's ability to efficiently process that input.

Although a body of research shows that processing measures in infancy and toddlerhood predict future vocabulary size, it is unclear whether this relationship holds in older children, even just beyond toddlerhood. Are children who are more efficient at processing language better word-learners more generally, or is this relationship observed only at the earliest stages of word learning? Children's vocabularies rapidly develop from about 18 months on, with 30-month-olds producing five times as many words as 18-month-olds. Furthermore, some of the early variability in vocabulary size disappears by preschool age (e.g., Paul, 1993; Rescorla, Mirak, & Singh, 2000). Failure for this trend to hold for preschool children would imply that processing is more critical for word learning at younger ages. Perhaps, preschool children have accumulated enough practice at recognizing familiar words that a processing advantage no longer translates into an advantage in learning words.

Additionally, it is not clear how environmental and child-level factors will interact in older children. Given the consistent positive relationship between language input measures and vocabulary, we would expect input quantity to predict vocabulary growth. In Newman et al. (2015), input repetitiveness and processing predicted vocabulary size independently, whereas in Weisleder and Fernald (2013) and Hurtado et al. (2008), the role of input worked indirectly, as mediated by lexical processing. In an older cohort, however, children might be fast enough at recognizing words such that input and processing independently predict word learning.

In this study, we examined the same kinds of variables as Weisleder and Fernald (2013): lexical processing, amount of language input, and vocabulary size at a future time. Because our study involved older children (28–39 months at Time 1), and we used different tasks. Specifically, we used direct measures of vocabulary at *both* time points, and we measured lexical processing using the more demanding four-image visual world paradigm rather than the two-image looking-while-listening paradigm. Because we have vocabulary measures at both time points, we can study how the environmental and child-level factors predict *change* in vocabulary, as opposed to future vocabulary size. Differences in word-learning trajectories emerge by 18 months (e.g., Frank, Braginsky, Yurovsky, & Marchman, 2016). Rowe, Raudenbush, and Goldin-Meadow (2012) found that the shape of a child's vocabulary growth-curve from 14 to 46 months predicted vocabulary size at 54 months, so it is important to factor in a child's prior vocabulary when studying how they accumulate new words.

Research Questions

We asked whether language input and lexical processing efficiency at age 2 ½–3 predicted vocabulary size one year later. At the heart of the matter is whether children who are more efficient at processing language are better word-learners. Processing measures in infancy and toddlerhood both predict future vocabulary size, and this study asks whether that effect still holds in this older cohort.

We also asked how input and processing interacted with each other. One possible interaction is a moderating relationship where processing efficiency constrains the effect of input on vocabulary growth. This outcome would imply that word recognition efficiency is a bottleneck for word learning, even at age 3. Alternatively, we might not observe any relationships between language input and processing. In this case, processing does not constrain a child's ability to learn from ambient speech, so that both support word learning independently.

Finally, we asked whether language input and processing predict differences in vocabulary growth. Specifically, we asked whether these measures are useful predictors of future vocabulary size when controlling for concurrent vocabulary size.

Methods and Measurements

Participants

We report all measurements and data exclusions following guidelines in Nosek et al. (2014). We examined data from the first two time points of a longitudinal study of preschoolers from English-speaking households. At Time 1 in the study, the children were 28–39 months old. During Time 1, we collected our measures of language input, lexical processing and vocabulary. At Time 2, we collected follow-up vocabulary measures when the children were 39–52 months old.

A total of 172 children provided vocabulary, processing and input data at Time 1. We excluded 5 children with cochlear implants from the present analysis. We also excluded 16 children identified by parents as late-talkers. Of the remaining children, 139 provided vocabulary measures at Time 2. As detailed below, we excluded 4 children for having inadequate home-language recordings and 26 children for having unreliable eyetracking data. A final total of 109 children were used in the vocabulary analyses. A small subset of the Time 1 vocabulary and eyetracking data (n = 14) was previously reported in Law, Mahr, Schneeberg, and Edwards (2016), which analyzed vocabulary size and concurrent lexical processing in a diverse group of participants. All children underwent a hearing screening at both time points, and they had normal speech, language, and cognitive development according to parent report.

Stimuli were presented in children's home dialect, either Mainstream American English (MAE) or African American English (AAE). We made an initial guess about what the home dialect was likely to be based on a number of factors, including the recruitment source and the child's address. For most children, the home dialect was MAE. If we thought the home dialect might be AAE, a native AAE speaker who was a fluent dialect-shifter was scheduled for the lab visit, and she confirmed the home dialect by listening to the caregiver interact with the child during the consent procedure at the beginning of the visit. AAE was the home dialect for 4 of the 109 participants.

Several other measurements were collected as part of the longitudinal study. They are not analyzed here because we limit attention to only the measures relevant for the analysis of input, processing, and vocabulary. Other unanalyzed Time 1 tasks were two pictureprompted word-repetition tasks (Edwards & Beckman, 2008), an eyetracking task with mispronunciations of familiar words (Law & Edwards, 2015), a minimal pair discrimination task (based on Baylis, Munson, & Moller, 2008), a verbal fluency task (WJ-III Retrieval Fluency subtest, Woodcock, McGrew, & Mather, 2001), a shape stroop task (Carlson, 2005), and an articulation test (GFTA-2, Goldman & Fristoe, 2000). Parents completed the *MacArthur-Bates Communicative Development Inventory* (Fenson et al., 2007), an inventory about executive function (BRIEF-P, Gioia, Espy, & Isquith, 2003), a survey about early literacy practices in the home (Senechal, 2006), and a demographic survey that included a multiple-choice question on maternal education level. A similar test battery was used at Time 2, with the addition of new tasks targeting phonological awareness (CTOPP-2, Wagner, Torgesen, Rashotte, & Pearson, 2013) and speech perception (SAILS task in Rvachew, 2006).

Vocabulary

At both time points, children received the *Expressive Vocabulary Test, 2nd Edition* (EVT-2, Williams, 2007) and its receptive counterpart, the *Peabody Picture Vocabulary Test, 4th Edition* (PPVT-4, L. M. Dunn & Dunn, 2007). In the expressive test, children were presented an image and had to name it. In the receptive test, children were presented four images and had to select a named image. For our analyses, we used growth scale values provided by each test; these values transform raw scores (words correct) into a scale that grows linearly with age.

Language input

Language input data was collected using a Language Environment Analysis (LENA) digital recorder, a small device worn by a child (Ford, Baer, Xu, Yapanel, & Gray, 2008). The device records all audible sounds for up to 16 hours. The recorder and instructions for using it were given to families. We instructed families to activate the recorder in the morning and record a typical day for the child. LENA software analyzed each recording to generate a summary of the child's language environment (Ford et al., 2008). The measures included 1) hourly word-counts of adult language in the child's environment, 2) hourly number of child-adult and adult-child conversational turns, 3) hourly proportions of meaningful (nearby) speech, distant speech, noise, television/electronics, and silence, and 4) hourly number of child vocalizations.

We computed the averages of each of these hourly measurements, excluding data from hours recorded after midnight. We computed the duration of the remaining before-midnight data in seconds, computing the number of hours from the number of seconds. This adjustment corrects for hours where the recording started midway through an hour. The average hourly adult word count then was the total adult word count in the recording divided by the number of hours. Our procedure differed that from Weisleder and Fernald (2013): That study only used the adult word counts from segments that coders had classified as child-directed.

We excluded recordings that might provide unreliable information. We excluded 3 recordings with less than 10 hours of data recorded before midnight, because such recordings undersampled the child's day. LENA software documentation also recommends that recordings be at least 10 hours in duration (LENA Foundation, 2015). We excluded 1 recording from a child who did not wear the device.

Lexical processing

Eyetracking procedure—To measure lexical processing, we used the visual world paradigm, an experimental procedure that has been used with children and adults (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; Huang & Snedeker, 2011; Law et al., 2016; McMurray, Samelson, Lee, & Tomblin, 2010). In this paradigm, images of objects are presented onscreen followed by a prompt to view one of the images. An eyetracker records the participant's gaze location over time. By examining how gaze changes in response to speech, we study the time course of word recognition. This particular experiment was described and analyzed in detail in Law et al. (2016).

In this experiment, four photographs of familiar objects appeared on a computer display. During a trial, a spoken prompt directed the child to view one of the images (e.g., *find the fly*). One image was the target (e.g., *fly*). The other distractor images contained a semantically related word (*bee*), a phonologically related word (*flag*), and an unrelated word (*pen*). Target words were presented in carrier frames (*see the* or *find the*). Children heard stimuli that matched their home dialect, either MAE or AAE. We recorded the stimuli from two young adult female speakers, one a native speaker of MAE and the other a native speaker of AAE. As noted above, 105 children came from families where MAE was spoken at home and received stimuli recorded in MAE; 4 children came from families where AAE was spoken at home and received stimuli recorded in AAE. In a cross-sectional study (Law et al., 2016) with an equal number of AAE- and MAE-speaking children (n = 30 per group), we did not observe differences between two dialect versions after controlling for child-level variables. Therefore, we combined data from both dialect versions in the analysis below.

Children saw 24 unique trials (each with different target words) in an experimental block. Each word served as the target word once per block. Two blocks of the experiment (each with different trial orderings and images) were administered. A Tobii T60XL eyetracker recorded the location of a child's gaze on the screen at rate of 60 Hz.

Presentation of carrier/target was gaze-contingent. After 2 s of familiarization time with the images in silence, the experiment paused to verify that the child's gaze was being tracked. After 300 ms of continuous gaze tracking, the trial advanced. Otherwise, if the gaze could not be verified after 10 s, the trial advanced. This step ensured that for nearly every trial, the gaze was being tracked before playing the carrier phrase, or in other words, that the child was ready to hear the carrier stimuli. An attention-getter or motivator phrase (e.g., *check it out!*) played 1 s after the end of the target word. Every six or seven trials, an animation played onscreen and the experiment briefly paused to allow examiners to reposition or coach the child to pay attention.

Data screening—We began with data from 151 children with Time 1 vocabulary scores, eyetracking data and home-language recordings. Data from 18 children had to be excluded because of a timing error in the experiment protocol that caused the reinforcer phrase to play too early after the target word. Before data screening, we performed *deblinking* by interpolating short windows of missing data (up to 150 ms) if the child fixated on the same image before and after a missing data window. We examined data quality in the 2-s window following the onset of the target. A trial was considered unreliable if at least 50% of the eyetracking data during the 2-s window was missing (offscreen). If at least 50% of trials in a block were unreliable, that block was excluded. We excluded 28 such blocks; 11 children had all their eyetracking trials excluded in this way. After block-level screening, we excluded an additional 712 unreliable trials. After screening, 4712 reliable trials remained from 122 children. Finally, we downsampled our data into 50-ms bins, reducing the eyetracking sampling rate from 60 Hz to 20 Hz. This procedure smoothed out high-frequency noise by pooling together data from adjacent frames.

Growth curve analysis—A common measure in eyetracking studies of word recognition is an *accuracy growth curve* (Mirman, 2014). We compute this growth curve by aggregating the number of looks to each image over trials and calculating the proportion of looks to the target image at each time sample. (We ignored offscreen looks or looks between the images when computing this proportion.) The growth curve measures how the probability of fixating on the target changes over time. Figure 1 depicts each participant's raw accuracy growth curve and the overall mean of the growth curves. On average, a child had a 25% chance of viewing the target image at the onset of the target word and the chance of looking to the image increased as the word unfolded and eventually plateaued after the word ended.

We used a mixed-effects logistic regression model to estimate the probability of fixating on the target image over time for each participant. We fit the model using the lme4 package (vers. 1.1.15; D. Bates, Mächler, Bolker, & Walker, 2015) in the R programming language (vers. 3.4.3). Although our vocabulary analyses use data from 109 participants, we used eyetracking data from 122 typically developing participants to fit the growth curve model so the data from the 13 additional participants would strengthen the model. See the Supplemental Appendix for detailed model results.

We modeled time using a cubic orthogonal polynomial. That is, our predictors were a constant term, (linear) time^1, (quadratic) time^2 and (cubic) time^3, and the time terms were scaled and centered so they were orthogonal and therefore uncorrelated. Because we used transformations of time, the constant did not estimate the predicted value at time = 0, but instead it estimated the area under the curve: the average log-odds of fixating on the target over the whole window.

The fixed effects of this model estimated an accuracy growth curve for an average participant. Of interest were the constant and linear-time terms. Because the constant term corresponded to the area under the growth curve, the model estimated an average probability of −0.297 logits (.43 proportion units) over all time samples. The linear time term captured the overall steepness of the growth curve. Ignoring the quadratic and cubic features of the growth curve, the linear term estimated an increase of 0.05 logits per 50 ms. At 0 logits (.5 proportion units), where the logistic function is steepest, an increase of 0.05 logits corresponds to an increase of .012 proportion units. At chance performance (.25 proportion units), this effect corresponds to an increase of .009 proportion units.

We allowed the constant and time terms to vary randomly within participants. These random effects quantified how an individual child's growth curve differed from the group average, so they provided measures of individual differences in lexical processing. Specifically, the constant terms provided a measure of overall accuracy, and the linear-time terms provided a measure of processing efficiency.

To visualize model-derived lexical processing measures, we divided the 109 children in the main analysis into thirds based on their linear-time coefficients. The faceted plot in Figure 2 shows growth curves for children with low, middle, and high linear trends, and the curves become steeper as the linear trend increases. For example, in the interval from 500 to 1,500 ms, each group's average proportion of looks to the familiar image increased by .17, .32 and .44. For children with higher slopes, the probability of fixating on the named image increases more quickly over time, so these children demonstrate more efficient lexical processing.

We can also quantify the lexical processing efficiency of each group by calculating the average linear-time parameter in each group and determining how much the probability increases when the average linear-time estimate is added to 0 logits. The predicted increase was .007 proportion units per 50 ms for children in the bottom group, .013 for children in the middle group, and .018 for children in the top group. By this measure, the children in the fastest group were more than twice as fast as the children in the bottom group.

Accuracy was related to processing efficiency. The by-child constant and linear time random effects were moderately correlated, $r = .31$; the children with steeper growth curves looked more to the target overall. The groups visualized had average looking proportions of .4, .45, and .45. Peak accuracy was also related to processing efficiency. We computed an asymptote for each child's growth curve as the median value from 1,500 to 2,000 ms, and the average asymptote for each group was .51, .63, and .71. These asymptotes were highly correlated with by-child constant effects, $r = .79$, and linear time effects, $r = .80$.

Analyses

Descriptive statistics

Table 1 presents summary statistics. The EVT-2 and PPVT-4 standard scores describe a child's ability relative to their age using an IQ-like scale (mean = 100, SD = 15). The children in this cohort had vocabulary scores approximately 1 SD greater than test-norm averages. Receptive and expressive vocabulary growth scale scores were highly correlated at both time points, $r_{T1} = .79$, $r_{T2} = .78$. Table 2 presents correlations for Time 1 measures. Most (92) of the children came from high maternal-education families (i.e., college or graduate degrees). Of the remaining children, 11 came from middle maternal-education families (at least two years of college, associate's degree, or trade school degree), and 6 from low maternal-education families (high school diploma or less, or less than two years of college).

Regression analyses

We used a Bayesian, estimation-based analytical approach: The aim is to estimate the magnitude and direction of effects as well as the uncertainty about those effects. In Bayesian models, we update our prior information based on the likelihood of the data—in other words, how well the data "fit" that prior information. The updated prior information constitutes the posterior distribution. Each sample from the posterior distribution represented a plausible set of parameters that is consistent with the observed data. We used this technique so that we could provide 95% uncertainty intervals for the parameter estimates. These intervals have an intuitive interpretation: We can be 95% certain the "true" value of a parameter, for the given model and data, is contained within its 95% uncertainty interval. This feature differs from frequentist confidence intervals which do not contain any distributional information about a given statistical effect (Kruschke & Liddell, 2017). Because the posterior contains plausible parameter values, we can measure our uncertainty about an effect's magnitude and direction. For example, if an interval spanned a large positive range, e.g., [3, 24] for some IQ-like standard scores, we would conclude that the effect is positive but that the size of the effect was very uncertain.

We fit Bayesian linear regression models using Stan (Carpenter et al., 2017) via the RStanARM package (vers. 2.17.3) in R. All predictors and outcome measures were scaled to have a mean of 0 and standard deviation of 1. We used weakly informative normal distributions as the priors of regression parameters: Intercept ~ *Normal*(μ = 0 [mean], σ = 5 [SD]) and Other Effects ~ *Normal*(μ = 0, σ = 1). This prior information implies that before seeing the data, we consider negative and positive effects to be equally plausible ($\mu = 0$), and we expect 95% of plausible effects to fall between ±1.96. We call this distribution "weakly informative" because of disciplinary expectations. In child language research, an effect where a 1-SD change in x predicts a 1-SD change in y represents a profound effect. Because our prior information generously includes such effects, they are "weakly informative".

Hamiltonian Monte Carlo sampling was performed on four chains each with 1,000 warm-up draws and 1,000 sampling draws, yielding 4,000 total draws from the posterior distribution. For all parameters reported, we used the median value of the parameter's posterior distribution as its "point" estimate. These median parameters values were used to calculate R^2 statistics as the conventional, unadjusted ratio of explained variance over total variance: R^2 = Var($\hat{y}$ [fitted]) / Var(y [observed]).

In all analyses, we used standardized average hourly adult word count as our measure of language input, and standardized linear-time coefficients (growth curve slopes) as our measure of processing efficiency. There was a small positive association between language input and lexical processing efficiency at Time 1, R^2 = .055, such that a 1-SD increase in input (an additional 468 words per hour) predicted a 0.24-SD increase in lexical processing efficiency, 95% Uncertainty Interval [0.06, 0.42].

Expressive vocabulary—There was a modest effect of input on expected expressive vocabulary at Time 2, R^2 = .074. A 1-SD increase in input predicted an increase of vocabulary of 0.27-SD units, UI [0.08, 0.45]. There was a strong effect of lexical processing, R^2 = .238. A 1-SD increase in processing efficiency predicted an increase in vocabulary of

0.49-SD units, UI [0.32, 0.66]. Estimates from each of these models are presented in Figure 3.

We also regressed expressive vocabulary onto input and processing, R^2 = .264. We observed a reliable effect of processing: For a child with an average amount of language input, a 1-SD increase in processing predicted an increase in vocabulary of 0.45 SD-units, UI [0.27, 0.61]. For a child with average processing efficiency, however, a 1-SD increase in input predicted a 0.17-SD increase in vocabulary size, UI [0.00, 0.34]. Lexical processing was a stronger predictor of vocabulary size than home language input, but there was also a modest, positive association between adult word counts and future expressive vocabulary size. There was not a credible Processing × Input interaction effect. That is, both positive and negative interaction effects were plausible, UI [−0.20, 0.13].

Receptive Vocabulary—There was a moderate effect of average hourly adult word count on receptive vocabulary size, R^2 = .107. A 1-SD increase in input (468 words per hour) predicted an increase of 0.33 SD units, 95% Uncertainty Interval [0.15, 0.50]. There was a strong effect of lexical processing efficiency, R^2 = .292. A 1-SD increase in processing efficiency predicted an increase in vocabulary of 0.54 SD units, UI [0.37, 0.70]. Estimates from each model are depicted in Figure 4.

We also regressed vocabulary onto input and processing efficiency. Both predictors were associated with vocabulary size, R^2 = .334. There was a strong effect of processing, β_{proc} = 0.49 SD units, UI [0.32, 0.66], whereas there was a modest effect of input, β_{input} = 0.21, UI [0.05, 0.37]. Because both input and processing showed positive effects, we also tested whether processing moderated the effect of input. There was not a credible Processing × Input interaction effect, UI [−0.15, 0.18]. These results indicate that lexical processing was a more robust predictor of future receptive vocabulary than average hourly adult word count, and also that adult word count had a positive effect on vocabulary over and above lexical processing ability.

Vocabulary Growth

We showed above that lexical processing efficiency and language exposure predicted vocabulary size one year later. These analyses are not adequate models of vocabulary *growth* because they do not account for vocabulary size at Time 1. If we think of home language input as a treatment variable—as language enrichment interventions do—then the analyses above ignored the pretreatment outcome levels.

The following analyses included Time 1 vocabulary size as a covariate so that we could model the effects of input and lexical processing. For each analysis, we started with a reference model in which we regressed vocabulary scores at Time 2 onto Time 1 vocabulary scores. We then added other predictors to see whether they had a credible effect *over and above* Time 1 vocabulary. These models allow us to examine the “value-added” properties of language exposure and lexical processing efficiency. The best performing model for each vocabulary type are described in detail in the Supplemental Appendix.

Expressive Vocabulary—As expected, there was a strong relationship between Time 1 and Time 2 expressive vocabularies, R^2 = .632. A 1-SD increase vocabulary scores at Time 1 predicted a 0.79-SD increase at Time 2, 95% Uncertainty Interval [0.68, 0.91]. There was not a credible effect of adult word count, UI [−0.12, 0.13]. There was no longer a 95% credible effect of processing, UI [−0.04, 0.24]. The posterior distribution of the processing effect was mostly positive, $P(0 < \beta_{\text{Processing}}) = .924$. If we stipulate that values between [0, 0.05] are so small that they are *practically equivalent* to 0, then 76.3% of posterior samples showed a non-null positive effect. Therefore, the data *suggests* a positive effect of lexical processing on expressive vocabulary growth. There was not a credible interaction between input and lexical processing efficiency, UI [−0.08, 0.15].

We compared these models and ones reported earlier using the Widely Applicable Information Criterion (WAIC; Table 3) computed via the loo R package (vers. 1.1.0; Vehtari, Gelman, & Gabry, 2017). Like other information criteria metrics (e.g., AIC or BIC), the WAIC estimates a model's predictive accuracy for out-of-sample data, and when comparing two models, the one with the lower WAIC is preferred. Because each observation independently contributes to the overall WAIC value, the WAIC is accompanied by a standard error (Vehtari et al., 2017) which helps quantify the uncertainty around WAIC point values. We also computed Akaike weights for WAIC values; these values provide a relative weighting or conditional probability estimate for each model (Wagenmakers & Farrell, 2004).

The models that do not include Time 1 vocabulary should be given no weight. Of the other models, we prefer the models without language input over those that include this predictor. Finally, we assign relatively equal weight to the model with just Time 1 vocabulary and the model with both lexical processing and Time 1 vocabulary. We would expect these models to perform the best on new data. Model comparison therefore provided little confirmatory support for a positive effect of lexical processing over and above Time 1 vocabulary.

Receptive Vocabulary—There was a strong relationship between Time 1 and Time 2 receptive vocabulary, R^2 = .584. A 1-SD increase in vocabulary at Time 1 predicted a 0.76-SD increase at Time 2, UI [0.64, 0.88]. There was a positive effect of adult word count over and above Time 1 vocabulary such that a 1-SD increase in input predicted a 0.15-SD increase in expected vocabulary, UI [0.03, 0.27], R^2 = .606. Similarly, a 1-SD increase in processing efficiency predicted an increase in receptive vocabulary of 0.23 SD units, UI [0.10, 0.37], R^2 = .626.

We also regressed receptive vocabulary onto all three predictors, R^2 = .640. There was a small effect of input over and above Time 1 vocabulary and lexical processing, $\beta_{\text{Input}} = 0.12$, UI [0.00, 0.24]. There was a moderate effect of processing, $\beta_{\text{Processing}} = 0.21$, UI [0.08, 0.35]. We did not observe a credible interaction of input and lexical processing, UI [−0.07, 0.16].

We compared the models using the WAIC (Table 4). We would expect the models with Time 1 vocabulary, lexical processing and language input as predictors to have to the best predictive accuracy on out-of-sample data. The most important variables for reducing WAIC were Time 1 vocabulary, followed by lexical processing, and lastly language input.

Receptive-Expressive Differences—Once we took Time 1 vocabulary into account, we observed different predictive effects of adult word count and lexical processing for expressive versus receptive vocabulary. For expressive vocabulary, input no longer had a credible effect, and lexical processing probably had a small positive effect but evidence for this effect is limited. In contrast, both predictors independently showed positive effects on receptive vocabulary, although the processing effect was larger than the input effect.

Based on these analyses alone, however, it would be invalid to claim receptive vocabulary was *more* sensitive to child-level factors than expressive vocabulary (Gelman & Stern, 2006). To evaluate these differences between receptive and expressive vocabulary, we have to estimate them. To compare both vocabulary outcomes simultaneously, we fit a multivariate regression model using Stan tools from the brms R package (vers. 2.1.0; Bürkner, 2017). As above, all variables were standardized to have mean 0 and standard deviation 1. We regressed Time 2 vocabulary onto Time 1 vocabulary, language input, lexical processing and the input-processing interaction for each vocabulary type as in preceding analyses. But to join the two outcomes, we also modeled the correlation between the residual error terms σ_{Rec} and σ_{Exp}. The error terms were moderately correlated, $\rho = .32$, UI [.13, .50].

The multivariate model maintained the results of the univariate models (see Figure 5): Strong effects of Time 1 vocabulary, reliable effects of input and processing on receptive vocabulary, and a suggestive effect of processing on expressive vocabulary. For each posterior sample, we computed the difference between receptive and expressive vocabulary coefficients (e.g., $\beta_{\text{Input[Diff]}} = \beta_{\text{Input[Rec]}} - \beta_{\text{Input[Exp]}}$), yielding a distribution of effect differences. Input had a stronger effect on receptive vocabulary that expressive vocabulary, $\beta_{\text{Input[Diff]}} = 0.12$, $P(0 < \beta_{\text{Input[Diff]}}) = .954$. A similar difference was observed for lexical processing, $\beta_{\text{Processing[Diff]}} = 0.11$, although it was slightly less probable the receptive effect was greater than the expressive effect, $P(0 < \beta_{\text{Processing[Diff]}}) = .922$. Lexical processing *probably* had a stronger effect on receptive than expressive vocabulary.

Discussion

We asked how lexical processing efficiency and home language input predicted vocabulary size one year later in a large sample of preschoolers. We measured lexical processing using an eyetracking experiment and model-derived estimates of how quickly on average a child’s gaze shifted to a named image. We measured language input using the average number of adult words per hour from LENA recordings, and we measured expressive and receptive vocabulary directly using standardized tests. We first tested how language input and lexical processing at age 3 predicted vocabulary size at age 4 without controlling for age-3 vocabulary levels. In these *baseline* analyses, both measures reliably and independently predicted vocabulary size. The processing effect was 2–2.5 times larger in magnitude than the input effect. Lexical processing and language input were weakly correlated, $r = .24$, and there were no credible interaction Processing × Input effects. These baseline analyses support the conclusions that lexical processing efficiency was a more important predictor of future vocabulary than language input and that word recognition efficiency did not constrain the beneficial effects of language exposure on future vocabulary size.

We next examined how input and lexical processing related to vocabulary *growth*, by controlling for age-3 vocabulary. The best predictors of Time 2 expressive vocabulary were Time 1 expressive vocabulary followed by lexical processing. The processing effect was less certain and smaller in magnitude compared to the robust effects observed in models of receptive vocabulary or in the baseline models that did not control for Time 1 vocabulary. Comparison of the expressive vocabulary models indicated that one should assign approximately equal weight to a model with both lexical processing and Time 1 vocabulary and a model with only Time 1 vocabulary. In contrast, the best predictors of Time 2 receptive vocabulary were Time 1 receptive vocabulary, followed by lexical processing and adult word count. Both processing and input provided additional predictive information over and above Time 1 vocabulary, and lexical processing had a larger effect than hourly adult word count. Finally, we estimated the differences in the effects on receptive versus expressive vocabulary, and the input effect was larger for receptive vocabulary while the processing effect was *probably* larger for receptive vocabulary.

The difference in results for expressive versus receptive vocabulary was unexpected, given the reliable correlation between the two measures, r = .81 at ages 2–5 (L. M. Dunn & Dunn, 2007, p. 60). Child who heard more words from their caregivers could understand more words one year later, but they could not necessarily produce more words. Why would language exposure be more related to receptive than expressive vocabulary? The differences for expressive and receptive vocabulary may simply reflect differences in the tests used to measure them. Across the entire course of the PPVT-4, the prompt remains the same (*show me X*). But over the course of the EVT-2, the prompts change from *what is this* to include prompts that demand metalinguistic knowledge (e.g., *tell me another word for X*). Thus, it may be the case that the PPVT-4 measures only receptive vocabulary, while the EVT-2 measures both expressive vocabulary and metalinguistic ability. Lexical processing efficiency and language input may be less related to metalinguistic ability than they are to vocabulary size.

Alternatively, the different results for expressive and receptive vocabulary may reflect the fact that recognition is easier than production. Being able to name an object—to activate the word's semantic representation and its phonological representation then carry out a motor plan—demonstrates a greater sign of mastery than being able to associate the word to an appropriate referent. The children who heard more words had more experience and exposure to words, giving them a broad base of shallow knowledge for word recognition. This interpretation would suggest that measures of input diversity (input types) would be even more predictive of future receptive vocabulary than simple quantity. Support for this interpretation also comes from Edwards et al. (2014). Using structural equation modeling, they found a direct relation between SES and receptive vocabulary, but only an indirect relation between SES and expressive vocabulary; the relation between SES and expressive vocabulary was mediated by receptive vocabulary.

A similar line of reasoning applies to the processing effect. We measured processing as response speed during a listening task, not a naming task. It captures a child's ability to activate a word's semantic representation in a timely manner. These demands are more clearly related to our receptive vocabulary task, whereas the expressive vocabulary task additionally required the child to talk. Nevertheless, we concluded that faster processing *probably* predicted larger expressive vocabularies. Approximately, 75% of the posterior samples indicated a positive, non-null effect. Naming an object still requires activation of a word's lexical representations, so it makes sense that the lexical processing efficiency would still matter for expressive vocabulary.

Throughout our analysis, we never observed a credible interaction effect between lexical processing and language input. Word recognition efficiency did not constrain the beneficial effects of language exposure on future receptive vocabulary size. One interpretation of these findings is that these children were *fast enough* at recognizing words that processing did not impose more of a bottleneck on vocabulary growth. Developmentally, that bottleneck may be observed in younger children than those in this sample. The youngest children in this study were 28 months, an age at which the average child produces about 500 words and recognizes at least 3 times that amount. In contrast, at 18 months the average child produces only about 50 words and recognizes about 250 (Frank et al., 2016). After about 18 months, children's vocabularies start increasing rapidly. At this point, it may be the case that processing efficiency no longer interacts with the quantity of language input. More research on children from 18 to 30 months is needed to evaluate this claim and to determine the time course of the relation among processing efficiency, language input, and vocabulary growth.

Our study elaborated on the work of Weisleder and Fernald (2013), but differed in important ways. Notably, our sample included children who were older in age, and we tested these children's vocabularies directly. Additionally, we used LENA's automated measures of adult word count, whereas that study used LENA word counts from just the segments of recordings that listeners had classified as child-directed. Limitations of our study include its observational design and the relatively homogeneous demographics of the families. This study was observational, so the analyses here describe statistical relationships. We did not manipulate language input, so we cannot establish causal links between language input and other measures. Moreover, most (92/109) children came from high maternal-education families (i.e., college or graduate degrees), whereas Weisleder and Fernald (2013) recruited 29 children from low-SES families. The combined SES and age differences make it difficult to compare these two studies directly. Although most of the participants in our study were demographically homogeneous, they varied in language input, vocabulary size, and processing efficiency, so they provided an informative test of how processing and input predict word learning. We found input and processing had positive effects on receptive vocabulary growth, but had little influence on expressive vocabulary growth.

Once we controlled for Time 1 vocabulary size, the effects of lexical processing and language input became less robust and less certain, especially for expressive vocabulary. It is essential that studies about changes in vocabulary size obtain a baseline vocabulary measurement. Our results were different when we measured predictors of vocabulary size as compared to predictors of vocabulary growth. This is not surprising, given that individual differences in vocabulary size are observed at 12 months or even earlier and increase with age. Unlike previous studies, this study measured both receptive and expressive vocabulary —if only one of these measures had been used as the sole measure of vocabulary knowledge, we would have drawn different conclusions.

Our findings have important implications for interventions aimed at increasing children's vocabulary. First and foremost, these interventions must start early. At Time 1, children in this study were only 28 to 39 months of age. Still, for both receptive and expressive vocabulary, vocabulary size at this young age was by far the strongest predictor of vocabulary size one year later. Based on our results, increasing the quantity of linguistic input for 3-year-olds is not going to be an effective intervention strategy. Other research has suggested that children attend to different features of their ambient language as their language abilities develop (cf. review in Lidz & Gagliardi, 2015; Rowe, 2012). For children of this age, sheer quantity of language input may not be as relevant or predictive as complexity or other features of the child-directed speech. The lack of evidence for quantity effects does not imply that quality is any more important or predictive of vocabulary growth at this age, but it suggests that measures like adult word count provide only a first approximation about the number of informative examples and learning opportunities available to the child. Language stimulation benchmarks and goals for children of this age, we would conclude, are better framed as time spent on activities, such as shared-book reading, or discussing events outside of the here and now (Rowe, 2012), rather than focusing simply on increasing language input.

These findings add to our understanding for vocabulary development. They suggest that the early relations among processing efficiency, input, and vocabulary size may decrease as all children become more efficient language processors in their third year of life. While differences in processing efficiency continue to exist, they may not create the bottleneck on vocabulary growth that is observed for children from 18 to 24 months. They also add to the very large literature supporting early intervention—children say their first words at about 12 to 14 months of age, but differences in vocabulary size even as early as 28 months are highly predictive of vocabulary growth one year later and are far more important than language input or processing efficiency.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIDCD R01 DC002932 to Jan Edwards, Mary E. Beckman, and Benjamin Munson; NICHD T32 HD049899 to Maryellen MacDonald; NIDCD T32 DC05359-10 to Susan Ellis Weismer; NIDCD Grant R01 DC012513 to Susan Ellis Weismer, Jan Edwards, and Jenny R. Saffran; and NICHD P30 HD003352 to the Waisman Center. We thank all the children who participated in this study and their families. We also acknowledge Mary E. Beckman, David Kaplan, Benjamin Munson, Franzo Law II, Alissa Schneeberg, Nancy Wermuth, and other members of the Learning To Talk labs at the University of Wisconsin–Madison and the University of Minnesota for their contributions to this research program.

References

Allopenna PD, Magnuson JS, Tanenhaus MK. Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. Journal of Memory and Language. 1998; 38(4):419–439. DOI: 10.1006/jmla.1997.2558

Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. Journal of Statistical Software. 2015; 67(1):1–48. DOI: 10.18637/jss.v067.i01

Baylis AL, Munson B, Moller KT. Factors affecting articulation skills in children with velocardiofacial syndrome and children with cleft palate or velopharyngeal dysfunction: A preliminary report. The Cleft Palate-Craniofacial Journal. 2008; 45(2):193–207. DOI: 10.1597/06-012.1 [PubMed: 18333642]

Bürkner PC. brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software. 2017; 80(1):1–28. DOI: 10.18637/jss.v080.i01

Carlson SM. Developmentally sensitive measures of executive function in preschool children. Developmental Neuropsychology. 2005; 28(2):595–616. DOI: 10.1207/s15326942dn2802_3 [PubMed: 16144429]

Carpenter B, Gelman A, Hoffman M, Lee D, Goodrich B, Betancourt M, Riddell Allen. Stan: A probabilistic programming language. Journal of Statistical Software. 2017; 76(1):1–32. DOI: 10.18637/jss.v076.i01

Cristia A, Seidl A, Junge C, Soderstrom M, Hagoort P. Predicting individual variation in language from infant speech perception measures. Child Development. 2014; 85(4):1330–1345. DOI: 10.1111/cdev.12193 [PubMed: 24320112]

Dunn, LM., Dunn, DM. Peabody Picture Vocabulary Test. Fourth. San Antonio, TX: Pearson Education; 2007.

Edwards JR, Beckman ME. Some cross-linguistic evidence for modulation of implicational universals by language-specific frequency effects in phonological development. Language Learning and Development. 2008; 4(2):122–156. DOI: 10.1080/15475440801922115 [PubMed: 19890438]

Edwards JR, Gross M, Chen J, MacDonald MC, Kaplan D, Brown M, Seidenberg MS. Dialect awareness and lexical comprehension of mainstream American English in African American English-speaking children. Journal of Speech, Language, and Hearing Research. 2014; 57(5): 1883–1895. DOI: 10.1044/2014_JSLHR-L-13-0228 [PubMed: 24949596]

Fenson, L., Marchman, VA., Thal, DJ., Dale, PS., Reznick, JS., Bates, E. MacArthur-Bates Communicative Development Inventories: User's guide and technical manual. 2nd. Baltimore, MD: Brookes; 2007.

Fernald A, Marchman VA. Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. Child Development. 2012; 83(1):203–222. DOI: 10.1111/j.1467-8624.2011.01692.x [PubMed: 22172209]

Ford, M., Baer, CT., Xu, D., Yapanel, U., Gray, S. The LENA™ language environment analysis system: Audio specifications of the DLP-0121 (Technical report LTR-032). Boulder, CO: LENA Foundation; 2008. Retrieved from http://www.lenafoundation.org/customer-resources/technical-reports/

Frank MC, Braginsky M, Yurovsky D, Marchman VA. Wordbank: An open repository for developmental vocabulary data. Journal of Child Language. 2016; doi: 10.1017/S0305000916000209 [PubMed: 27189114]

Gelman A, Stern H. The difference between "significant" and "not significant" is not itself statistically significant. The American Statistician. 2006; 60(4):328–331. DOI: 10.1198/000313006X152649

Gioia, GA., Espy, KA., Isquith, PK. Behavior Rating Inventory of Executive Function - Preschool Version. Odessa, FL: Psychological Assessment Resources; 2003.

Goldman, R., Fristoe, M. Goldman-Fristoe Test of Articulation. Second. Minneapolis, MN: Pearson; 2000.

Hart, B., Risley, TR. Meaningful differences in the everyday experience of young American children. Baltimore: P.H Brookes; 1995.

Hoff E. The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. Child Development. 2003; 74(5):1368–1378. DOI: 10.1111/1467-8624.00612 [PubMed: 14552403]

Hoff E. How social contexts support and shape language development. Developmental Review. 2006; 26(1):55–88. DOI: 10.1016/j.dr.2005.11.002

Huang YT, Snedeker J. Cascading activation across levels of representation in children's lexical processing. Journal of Child Language. 2011; 38(03):644–661. DOI: 10.1017/S0305000910000206 [PubMed: 20738890]

Hurtado N, Marchman VA, Fernald A. Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. Developmental Science. 2008; 11(6):F31–9. DOI: 10.1111/j.1467-7687.2008.00768.x [PubMed: 19046145]

Huttenlocher J, Haight W, Bryk A, Seltzer M, Lyons T. Early vocabulary growth: Relation to language input and gender. Developmental Psychology. 1991; 27(2):236–248. DOI: 10.1037/0012-1649.27.2.236

Junge C, Cutler A. Early word recognition and later language skills. Brain Sciences. 2014; 4(4):532–559. DOI: 10.3390/brainsci4040532 [PubMed: 25347057]

Kruschke JK, Liddell TM. The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. Psychonomic Bulletin & Review. 2017; :1–29. DOI: 10.3758/s13423-016-1221-4

Kuhl PK, Conboy BT, Coffey-Corina S, Padden D, Rivera-Gaxiola M, Nelson T. Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). Philosophical Transactions of the Royal Society B: Biological Sciences. 2008; 363(1493):979–1000. DOI: 10.1098/rstb.2007.2154

Lany J. Lexical-processing efficiency leverages novel word learning in infants and toddlers. Developmental Science. 2017; :e12569.doi: 10.1111/desc.12569 [PubMed: 28597549]

Law F II, Edwards JR. Effects of vocabulary size on online lexical processing by preschoolers. Language Learning and Development. 2015; 11(4):331–355. DOI: 10.1080/15475441.2014.961066 [PubMed: 26508903]

Law F II, Mahr T, Schneeberg A, Edwards JR. Vocabulary size and auditory word recognition in preschool children. Applied Psycholinguistics. 2016; doi: 10.1017/S0142716416000126 [PubMed: 28439144]

LENA Foundation. LENA Pro user guide. 2015. Retrieved from https://shop.lena.org/pages/download-center

Lidz J, Gagliardi A. How nature meets nurture: Universal grammar and statistical learning. Annual Review of Linguistics. 2015; 1(1):333–353. DOI: 10.1146/annurev-linguist-030514-125236

Marchman VA, Bates E. Continuity in lexical and morphological development: A test of the critical mass hypothesis. Journal of Child Language. 1994; 21(02):339–366. DOI: 10.1017/S0305000900009302 [PubMed: 7929685]

Marchman VA, Fernald A. Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. Developmental Science. 2008; 11(3):F9–16. DOI: 10.1111/j.1467-7687.2008.00671.x [PubMed: 18466367]

McMurray B, Samelson VM, Lee SH, Tomblin JB. Individual differences in online spoken word recognition: Implications for SLI. Cognitive Psychology. 2010; 60(1):1–39. DOI: 10.1016/j.cogpsych.2009.06.003 [PubMed: 19836014]

Mirman, D. Growth curve analysis and visualization using R. Boca Raton, FL: Chapman & Hall/CRC; 2014.

Newman RS, Rowe ML, Bernstein Ratner N. Input and uptake at 7 months predicts toddler vocabulary: The role of child-directed speech and infant processing skills in language development. Journal of Child Language. 2015; :1–16. DOI: 10.1017/S0305000915000446

Nosek, BA., Simonsohn, U., Moore, DA., Nelson, LD., Simmons, JP., Sallans, A., LeBel, EP. Standard reviewer statement for disclosure of sample, conditions, measures, and exclusions. Open Science Framework. 2014 February. Retrieved from https://osf.io/hadz3

Paul R. Patterns of development in late talkers: Preschool years. Journal of Childhood Communication Disorders. 1993; 15(1):7–14. DOI: 10.1177/152574019301500103

Rescorla L. Age 17 language and reading outcomes in late-talking toddlers: Support for a dimensional perspective on language delay. Journal of Speech, Language, and Hearing Research. 2009; 52:16–30. DOI: 10.1044/1092-4388(2008/07-0171) [PubMed: 18723598]

Rescorla L, Mirak J, Singh L. Vocabulary growth in late talkers: Lexical development from 2;0 to 3;0. Journal of Child Language. 2000; 27(2):293–311. [PubMed: 10967889]

Rowe ML. A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. Child Development. 2012; 83(5):1762–1774. DOI: 10.1111/j.1467-8624.2012.01805.x [PubMed: 22716950]

Rowe ML, Raudenbush SW, Goldin-Meadow S. The pace of vocabulary growth helps predict later vocabulary skill. Child Development. 2012; 83(2):508–525. DOI: 10.1111/j.1467-8624.2011.01710.x [PubMed: 22235920]

Rvachew S. Longitudinal predictors of implicit phonological awareness skills. American Journal of Speech-Language Pathology. 2006; 15(2):165–176. DOI: 10.1044/1058-0360(2006/016) [PubMed: 16782688]

Scarborough, HS. Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In: Fletcher-Campbell, F.Soler, J., Reid, G., editors. Approaching difficulties in literacy development: Assessment, pedagogy, and programmes. London: SAGE; 2009. p. 23-39.

Senechal M. Testing the home literacy model: Parent involvement in kindergarten is differentially related to grade 4 reading comprehension, fluency, spelling, and reading for pleasure. Scientific Studies of Reading. 2006; 10(1):59–87. DOI: 10.1207/s1532799xssr1001_4

Shankar, M. Empowering our children by bridging the word gap. Blog. 2014 June. Retrieved from https://www.whitehouse.gov/blog/2014/06/25/empowering-our-children-bridging-word-gap

Tsao FM, Liu HM, Kuhl PK. Speech perception in infancy predicts language development in the second year of life: A longitudinal study. Child Development. 2004; 75(4):1067–1084. DOI: 10.1111/j.1467-8624.2004.00726.x [PubMed: 15260865]

Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing. 2017; 27(5):1413–1432. DOI: 10.1007/s11222-016-9696-4

Wagenmakers EJ, Farrell S. AIC model selection using Akaike weights. Psychonomic Bulletin & Review. 2004; 11(1):192–196. DOI: 10.3758/BF03206482

Wagner, R., Torgesen, J., Rashotte, C., Pearson, NA. Comprehensive Test of Phonological Processing. Second. Austin, TX: PRO-ED; 2013.

Walley AC, Metsala JL, Garlock VM. Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. Reading and Writing. 2003; 16(1):5–20. DOI: 10.1023/A:1021789804977

Weisleder A, Fernald A. Talking to children matters: Early language experience strengthens processing and builds vocabulary. Psychological Science. 2013; 24(11):2143–52. DOI: 10.1177/0956797613488145 [PubMed: 24022649]

Williams, KT. Expressive Vocabulary Test. Second. San Antonio, TX: Pearson Education; 2007.

Woodcock, RW., McGrew, KS., Mather, N. Woodcock-Johnson III. Itasca, IL: Riverside Publishing; 2001.

Research highlights

- We examine how home language input and word recognition efficiency at 28–39 months predict vocabulary size one year later.
- Word recognition efficiency does not constrain the effect of home language input on word learning at this age.
- Receptive vocabulary is more sensitive to variability in language input and lexical processing than expressive vocabulary.

Author Manuscript

Author Manuscript

Author Manuscript

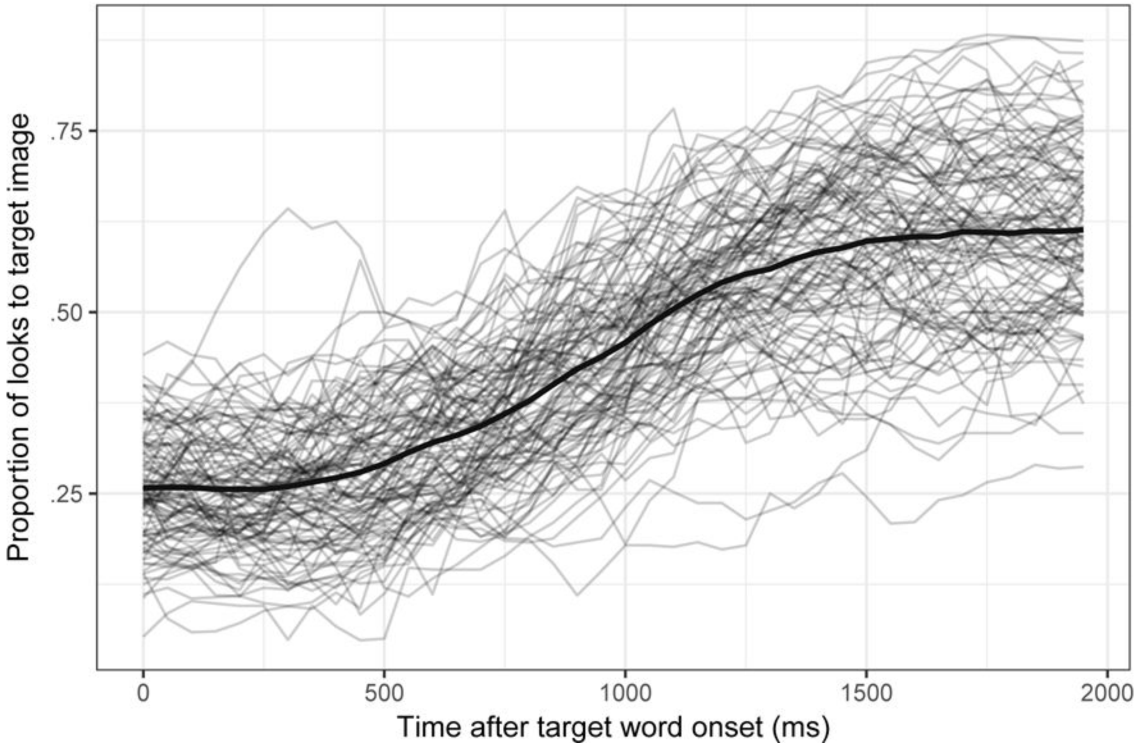

Figure 1.
Spaghetti plot of raw individual accuracy growth curves for 109 participants. Each light line represents the observed proportion of looks to the target image over time for one participant. The darker line represents the average of the growth curves.

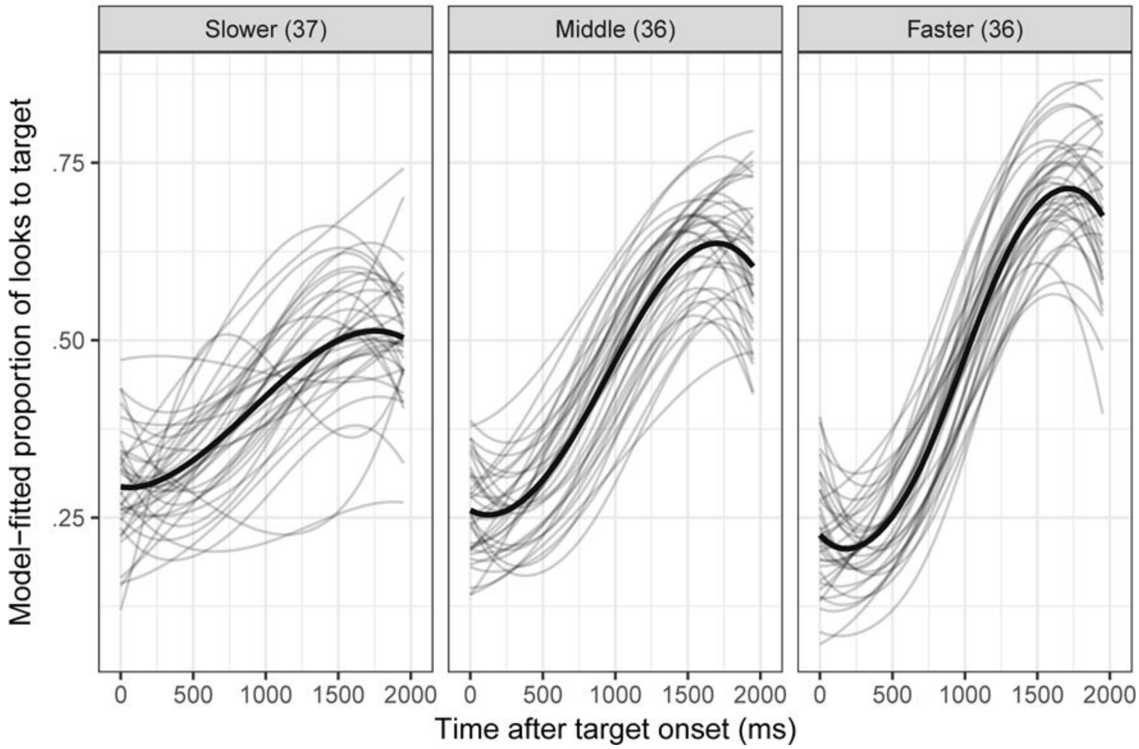

Figure 2.

Model-fitted accuracy growth curves for participants grouped by linear-time coefficients. Participants were divided into tertiles. Light lines represent model-estimated growth curves for individual children and dark lines represent the average of growth curves within each facet.

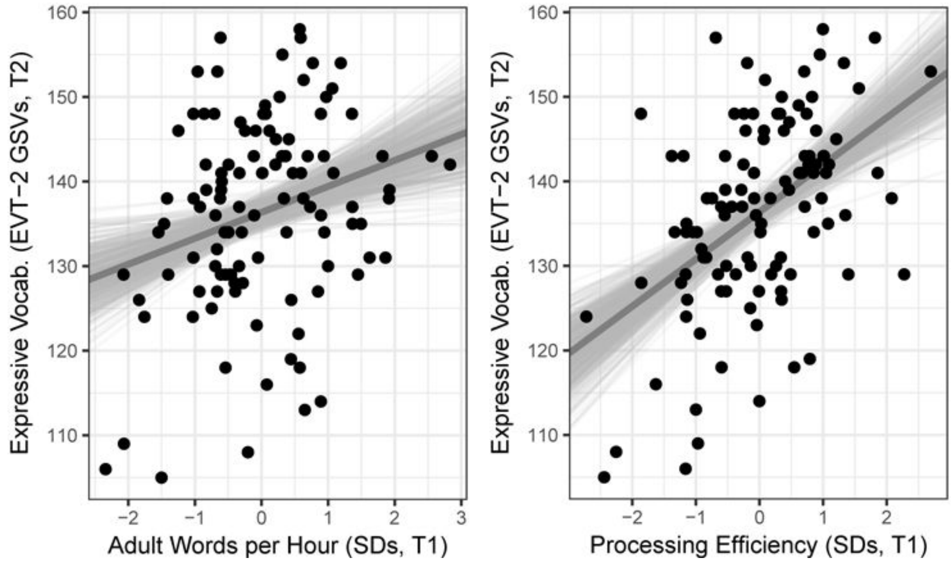

Figure 3.

Regression models for expressive vocabulary. The heavy line in each plot represents the median of the posterior distribution of the model. Light lines represent 500 random draws from the posterior. The lines are included to depict uncertainty of the modeled relationship.

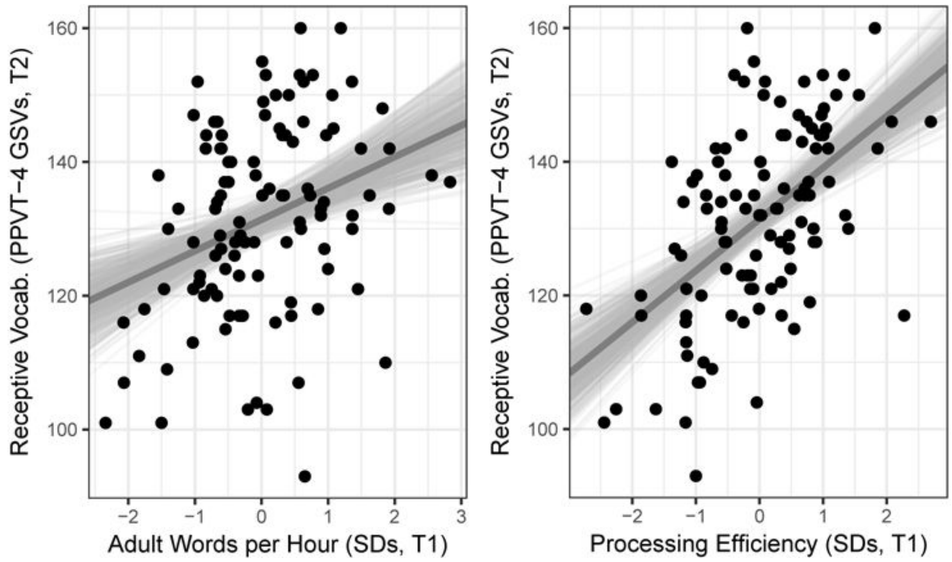

Figure 4.
Regression models for receptive vocabulary. The heavy line in each plot represents the median of the posterior distribution of the model. Light lines represent 500 random draws from the posterior. The lines are included to depict uncertainty of the modeled relationship.

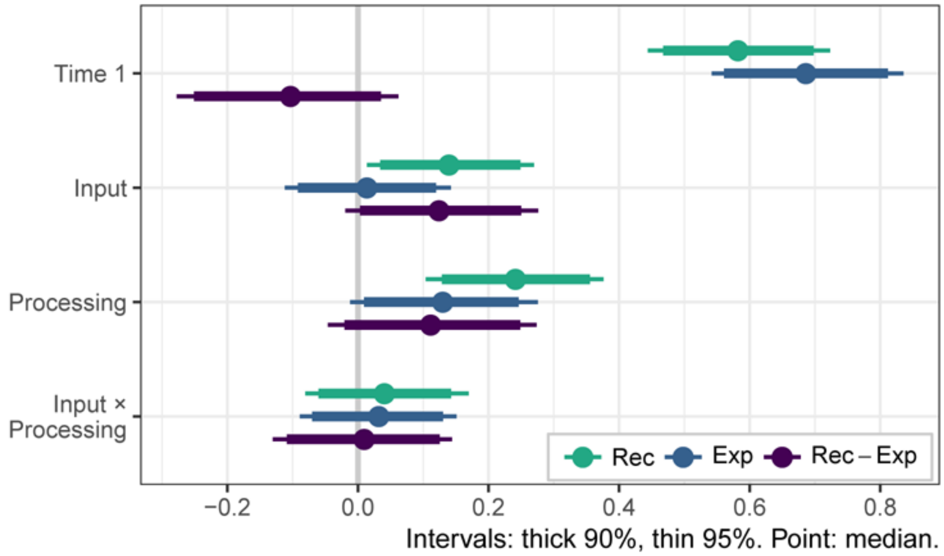

Figure 5.
Posterior median and 95% and 90% uncertainty intervals for the vocabulary effects and effect differences.

Table 1

Summary statistics for Time 1 and Time 2 (*N* = 109).

Time	Measure	Mean	SD	Range
1	Age (months)	32.9	3.4	28–39
	Hourly Adult Word Count	1207.0	467.6	112–2531
	Exp. Vocab. (EVT-2 GSVs)	118.9	11.6	81–148
	Exp. Vocab. (EVT-2 Standard)	118.4	14.2	81–160
	Rec. Vocab. (PPVT-4 GSVs)	109.2	16.6	70–151
	Rec. Vocab. (PPVT-4 Standard)	116.0	15.6	84–153
2	Age (months)	45.1	3.5	39–52
	Exp. Vocab. (EVT-2 GSVs)	136.4	11.4	105–158
	Exp. Vocab. (EVT-2 Standard)	119.8	14.6	80–155
	Rec. Vocab. (PPVT-4 GSVs)	131.4	14.3	93–160
	Rec. Vocab. (PPVT-4 Standard)	121.8	13.9	90–151

Table 2

Correlations among Time 1 measurements.

	Age (months)	EVT-2 GSVs	PPVT-4 GSVs	Processing efficiency
EVT-2 GSVs	.40			
PPVT-4 GSVs	.50	.79		
Processing efficiency	.31	.53	.47	
Hourly adult word count	−.12	.34	.24	.24

Table 3

Model comparisons for expressive vocabulary.

Predictors	WAIC ± SE	Akaike Weight
T1	205.1 ± 14.5	.388
T1 + Processing	205.6 ± 14.9	.301
T1 + Input	207.2 ± 14.4	.138
T1 + Input + Processing	207.3 ± 14.8	.127
T1 + Input + Processing + (Input × Processing)	209.4 ± 14.8	.046
Input + Processing	283.3 ± 14.6	.000

Table 4

Model comparisons for receptive vocabulary.

Predictors	WAIC ± SE	Akaike Weight
T1 + Input + Processing	206.8 ± 15.7	.516
T1 + Input + Processing + (Input × Processing)	207.9 ± 15.2	.303
T1 + Processing	209.1 ± 14.7	.170
T1 + Input	214.8 ± 15.7	.009
T1	218.8 ± 14.8	.001
Input + Processing	272.0 ± 14.9	.000