Running Head: Lexical Processing Efficiency and Vocabulary Size

Infants' Lexical Processing Efficiency is Related to Vocabulary Size by One Year of Age

Jill Lany
University of Notre Dame

Michael Giglio
University of Texas-Houston McGovern Medical School

Madeleine Oswald
The University of Chicago

Address Correspondence to:

Jill Lany, PhD
118C Haggar Hall
Department of Psychology
University of Notre Dame
Notre Dame, IN 46656
Email: jlany@nd.edu

Acknowledgements: This work was supported by funds from NSF BCS-1352443 and by Notre Dame's Institute for Scholarship in the Liberal Arts to J.L. The authors have no conflicts of interest to declare.

Abstract

By 15-18 months, infants' skill in interpreting familiar words, or lexical-processing efficiency (LPE) improves substantially, and is correlated with vocabulary size concurrently and several months later. Prior to this age, LPE is quite poor, and to date there is little evidence that it is related to vocabulary size. If this relation only emerges once infants have relatively good LPE, and also know a substantial number of words, it could suggest that the processes that support the rapid growth in vocabulary commonly observed as infants approach age 2 may not yet be functional in the earlier stages of lexical development. However, by using a modified LPE task we found that 12-month-olds with better LPE understood more words at that age, and also produced more words several months later. Thus, meaningful individual differences in LPE are already emerging by 12 months, and may support lexical development across the second year.

Adults are highly capable of interpreting speech from their native language, using even just the first phonemes of a word to find its referent (Marslen-Wilson & Zwisterlood, 1989). Infants typically find the visual referents of familiar spoken words much more slowly than adults, if at all, but their lexical recognition improves substantially between 15 and 24 months of age (Fernald, Perfors, & Marchman, 2006; Fernald, Pinto, Swingley, Weinbergy, & McRoberts, 1998). Here we refer to skill in lexical recognition as lexical-processing efficiency (LPE).

The gains that infants make in their LPE across the first several years are important for at least three reasons. First, they reflect deepening knowledge of the early-learned words and sentence structures that are commonly used in LPE tasks. Second, gains in LPE allow infants to keep up with speech in the moment. Third, and most relevant to the current work, beyond reflecting how well infants know and recognize familiar words, LPE appears to promote lexical development. Specifically, by the time they reach 18 months of age, infants with relatively good LPE already have larger vocabularies according to the MacArthur Communicative Development Inventory (MCDI), a widely used parent-report measure of vocabulary size, and also learn more words across the following 6 to 12 months, (Fernald & Marchman, 2012; Fernald, Marchman, & Weisleder, 2013; Weisleder & Fernald, 2013). Furthermore, at 16- to 18-months of age infants who are relatively fast at recognizing familiar words in a LPE task are also better able to learn novel nonce words than infants who are relatively slow, even when given the same amount of exposure to the new words (Lany, 2017). These data suggest that infants with relatively good LPE are better able to form robust and accurate lexical representations as they encounter new words, thereby supporting growth in the lexicon.

It is likely that a bidirectional synergy holds between gains in LPE and in vocabulary size, such that they support each other. Learning words over multiple exposures provides opportunities to practice encoding and recognizing word forms, as well as accessing their meanings, and thus the process of learning words is likely to hone the skills used for lexical recognition. Likewise, novel words can be easier to learn when they are surrounded by known words (e.g., Fisher, Gleitman, & Gleitman, 1991), and infants with better LPE are likely to be better able to encode such informative contexts. Furthermore, 18-month-olds are faster to recognize words that come from relatively dense semantic networks (Borovsky, Ellis, Evans, & Elman, 2016), suggesting that adding words to early lexical networks may facilitate lexical access.

Importantly, these findings suggest that a synergy between LPE and word learning may not be present in the very early stages of lexical development, emerging only after infants have formed relatively large, dense lexical networks. In fact, while there is substantial evidence that LPE and vocabulary size and growth are related in infants about 18 months and older, there is little evidence that LPE is related to vocabulary size in younger infants (Bergelson & Swingley, 2012; 2013; 2015; Reznick, 1990; Zangl, Klarman, Thal, Fernald, & Bates, 2005). For example, even though several studies have shown that infants recognize common words as early as 6 months of age, there little evidence that recognition skill on such tasks and vocabulary size are correlated before 14-16 months. Furthermore, aspects of younger infants' performance on LPE tasks suggest that important changes in lexical development may be happening around 15 months. Specifically, there are nonlinear improvements in infants' recognition performance at 14-16 months of age (Bergelson & Swingley, 2012; 2013; 2015). Given that the relation between LPE and vocabulary size and growth begins to hold more consistently at that age, it is possible that the synergy between them represents a new process contributing to lexical development.

However, the current evidence on relations between LPE and vocabulary size in infants prior to 15-18 months is not conclusive. The presence of a large improvement in recognition skills at around 15 months does not necessarily suggest that there is a discontinuity in how LPE relates to word learning, or to lexical development more broadly. Furthermore, the lack of evidence for significant relations between LPE and vocabulary size should be interpreted with caution, as most of the studies testing lexical recognition in younger infants were not primarily designed to test LPE-vocabulary size relations. Thus, the goal of the current work was to test whether the synergy between LPE and vocabulary size, which appears to reflect a key process supporting lexical development by 18 months, holds in younger infants. To that end, we designed a measure of LPE that we believed should be sensitive to individual differences at 12 months of age, and tested whether it is related to infants' vocabulary size concurrently and several months later. If LPE and vocabulary size are linked at 12 months, it would suggest that the synergy between LPE and vocabulary size supports early lexical development. Before describing the design of the current study, we more closely consider the existing work on LPE in younger infants.

Previous Findings on LPE and Vocabulary Size in Infants Younger than 15 Months

A handful of studies have investigated LPE in infants younger than 15-18 months (Bergelson & Swingley, 2012; 2013; Reznick, 1990; Swingley & Aslin, 2002; Zangl et al., 2005). In these studies, LPE was tested by presenting pictures of two objects (e.g., a baby and a dog) and tracking infants' gaze as the label for one was spoken (e.g., “Find the baby!”). The relative amount of looking to the target during the seconds after the label is presented, or Accuracy, has generally been used to assess comprehension or recognition skill in infants at this age. In some of these studies, evidence for successful comprehension was weak to nonexistent. For example, Reznick (1990) tested whether 8-, 14-, and 20-month-old infants showed evidence of comprehending 4 words; *butterfly*, *dog*, *infant*, and *woman*. Even when using a very lenient criterion (a 5% increase in looking to the target object after it was labeled), 8-month-olds showed evidence of comprehending, on average, 1 of the 4 words, and 14-month olds did not perform much better. This may not be surprising given that only the word "dog" is likely to have been known by infants at the tested ages. Zangl et al. (2005) tested comprehension in 12- to 31-month-old infants on a set of 24 words that are typically learned in the first 3 years, though some words (e.g., "horse" and "phone") are unlikely to have been familiar to the youngest infants tested. While older infants showed evidence of lexical recognition, it was not clear that the 12- to 14-month-olds reliably identified any of the referents (c.f. Fernald, Zangl, Portillo, & Marchman, 2008). Critically, in neither study was infants' LPE related to vocabulary size, as measured by the MCDI, before approximately 15-18 months.

In the studies reviewed above, the lack of evidence for a relation between LPE and vocabulary size could easily be because the tasks used to assess LPE were too difficult to reveal individual differences in the younger infants. However, even in studies in which infants showed better evidence of comprehension there is little evidence that LPE is related to vocabulary size. For example, Swingley and Aslin (2002) found that 14-month-olds recognized common words like “dog” and “shoe” when they were correctly pronounced, and when they were mispronounced, as evidenced by greater looking to the target picture than to the foil. However, infants’ recognition skill was unrelated to the number of words they knew according to the MCDI.

In a series of studies, Bergelson and Swingley found evidence of lexical recognition in infants as young as 6 months of age. In their initial study, they tested lexical recognition in 6- to 16-month-old infants using words for common foods and body parts, such as "milk" and "feet" (Bergelson & Swingley, 2012). They found that even the youngest infants showed some evidence of comprehension, though by 14-16-months performance was much better. Critically, MCDI measures of vocabulary size were obtained for infants 8 months and older, but they were uncorrelated with LPE at all ages (personal communication). In a subsequent study Bergelson and Swingley (2013) found evidence that by 10 months infants understood some words referring to actions, routines, such as “kiss” and “bye”, though not earlier. Recognition improved substantially by 14-16 months, and only at this older age was LPE correlated with vocabulary size. In an extension of these studies, Bergelson and Swingley (2015) replicated their findings that infants can comprehend common nouns and verbs by 6-9 months, and that infants whose parents reported them to say at least one word performed better than those who were not yet saying any words. Interestingly, they again found a sharp increase in comprehension as infants approached 15 months of age.

In sum, across a range of studies testing lexical recognition in infants younger than 15-18 months, there is consistent evidence of lexical recognition by 6 months of age, and that recognition improves noticeably at 14-16 months.

However, there is little evidence that LPE is related to parent-report measures of vocabulary size at these ages.

When and How do LPE and Vocabulary Size Become Related?

It is clear that LPE and vocabulary size are related by 18 months (Fernald et al., 2006; Fernald & Marchman, 2012). There is also evidence that they continue to be related across early childhood (Law & Edwards, 2014). At present, however, there is little evidence that LPE and vocabulary size are related prior to 15 months. The improvements in lexical recognition observed at 14-16 months may suggest the processes of word-learning change in important ways at this time (Bergelson & Swingley, 2015). Given the evidence reviewed above, it is possible that the relation between LPE and vocabulary size emerges at this age, representing a new mechanism of lexical development.

However, it is possible that a synergy between LPE and lexical growth is present but hard to detect in existing studies, which were not designed to capture such a relation. While infants showed evidence of comprehension well before their first birthdays in these tasks, a large set of words that potentially varied in familiarity was used in the assessments. If fluctuations in performance on trials testing relatively unfamiliar words reflects noise, rather than real differences in recognition, they may have obscured more meaningful differences in how well individual infants performed on more familiar words. Furthermore the parent-report measure of receptive vocabulary size may be noisy for very young infants, which could also mask an association with LPE.

Moreover, during the first year of life, infants make strides in skills that are likely to support real-time comprehension. For example, they become attuned to native language phonetics, refining their sensitivity to differences between speech sounds that are relevant in their native-language (Kuhl et al, 1995; Werker & Tees, 1984). They also become more skilled at identifying word forms in fluent speech (Jusczyk & Aslin, 1995) and recognizing a given word form when produced by different individuals (Houston & Jusczyk, 2000), or with different affect (Singh et al., 2004). Importantly, individual differences in the development of speech-sound discrimination and word-form recognition in the first year predict parent report measures of vocabulary size at age 2 (Tsao, Liu, & Kuhl, 2004; Singh et al., 2012; Junge et al., 2012). Skill in word-form recognition is likely to contribute to the ability to comprehend familiar word forms, and locate their referents quickly. Thus, these findings suggest that there are likely to be individual differences in infants' lexical recognition ability prior to amassing a large vocabulary, and that these individual differences may be related to word learning skill.

Given the ambiguity about the origins of individual differences in LPE and its relation to early lexical development, in the current study we wanted to conduct a strong test of whether individual differences in LPE are related to vocabulary size in infants at 12 months of age. To that end, we adapted the lexical recognition tasks used in previous research with the goal of increasing sensitivity to individual differences in infants' LPE. In particular, we tested infants on words that are likely to be known by most 12-month-olds (Easy words), as well as words less likely to be familiar to them (Hard words), and assessed LPE on each trial type separately. Following the design of previous work (Reznick, 1990; Bergelson & Swingley, 2012; 2013, 2015), we tested whether there are concurrent relations between performance on this LPE task and MCDI measures of vocabulary size. We also tested whether LPE at 12 months predicts vocabulary size several months later, as it does in older infants, and whether individual differences in performance on the LPE task are stable across development.

Methods

Participants

Participants were 66 infants who were between 12 months and 12 months and 30 days at the start of the study (i.e., Visit 1). Of those, 39 were female. Parental report confirmed that infants were born at full term and had no significant history of ear infections or developmental disorders. Infants were primarily Caucasian (there were 3 African American infants and 1 Hispanic infant) and from families with high levels of maternal education (4% had a high-school degree, 21% attended some college, and 75% had a college degree or higher). At Visit 1, infants' mean percentile scores for the receptive vocabulary size measure, assessed via parent report on the MCDI, was 43 with a range of 5 to 95. Infants whose parents indicated that they were exposed to a language other than English for more than 15 hours a week were not considered eligible. An additional 5 infants were tested but their data were not included because they had a history of chronic ear infections within the last year (n = 2), equipment failure (n = 2), and fussiness (n = 1).

Of the original sample, 48 infants (23 female) signed up for another study taking place in the lab when they were between 15 and 19 months of age (i.e., Visit 2). We assessed their LPE when they participated in those studies, allowing us to get a second measure of their real-time processing skills at an age when this task has yielded meaningful individual differences in the LPE measures. The data from 8 of these infants were excluded because they contributed insufficient usable trials to compute any measures of LPE (n = 3), because of fussiness (n = 2), parent interference (n = 1), falling asleep during testing (n = 1), and for a MCDI score more than 3 standard deviations away from the sample mean (n = 1). Thus 40 infants contributed LPE data to the Visit 2 measures.

The present study was conducted according to guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a parent or guardian for each child before any assessment or data collection. All procedures involving human subjects in this study were approved by the Institutional Review Board at the University of Notre Dame.

Materials and Procedure

Visit 1 Measures

LPE Task. Infants were tested on an LPE task using the Looking-While-Listening Procedure (see Fernald et al., 2008 for a general description). Eight words were selected for inclusion based on the likelihood that they would be familiar to infants according to MCDI norming data (Dale & Fenson, 1996). The estimates of the percentage of 12-month-olds likely to know the words were obtained from the original norming study (Dale & Fenson, 1996), but the results from Word Bank (Frank, Braginsky, Yurovsky, & Marchman, 2016) were nearly identical. The Easy words were likely to be known by a relatively high percentage of 12-month olds, and the Hard words were likely to be known by relatively few of them. Following the general design of Bergelson and Swingley (2012) we chose words from two categories, animals and foods, with referents that are easily recognizable in still pictures. For Easy words, we picked two words within each category that were estimated to be known by more than 50% of 12-month-olds, with the goal of roughly matching them on that metric. The Easy words were "doggie" (68.8% of 12-month-olds are reported to understand this word), "kitty" (49.7% comprehension rate), "milk" (58.6% comprehension rate), and "banana" (53.5% comprehension rate). We chose Hard words from the food and animal categories such that fewer than half of infants at 12 months were likely to know them. These words were "bear" (25.5% comprehension rate), "horse" (17.2% comprehension rate), "bread" (23.6% comprehension rate), and "apple" (17.2%

comprehension rate). Following Bergelson & Swingley, (2012), these items were presented in yoked pairs, such that a given food item was always presented with a given animal. The Easy word pairs were ‘doggie’ - ‘milk’, and ‘kitty’ - ‘banana’, and the Hard word pairs were ‘bear’ - ‘apple’ and ‘horse’ - ‘bread'). To heighten the visual interest of the materials, two different images were used for each object across trials.

At the start of each trial, the two images from a yoked pair appeared simultaneously on a 60” LCD screen, one in each of the bottom corners. After a 2-second silence, the target word was presented in one of two different sentence frames, e.g., "Find the kitty" or "Where's the kitty". The pictures remained on the screen for approximately 4 additional seconds, giving infants an opportunity to find the referent. Several tokens of each sentence, spoken by a native female English-speaker in an animated voice, were recorded. The best token of each was selected and edited to keep volume consistent across the trials. Each word was presented 4 times, 2 times in each of the 2 frames. After every fourth trial, an “attention getter" trial consisting of a colorful image paired with music, was presented to keep infants engaged in the task.

The presentation of items in the test trials counterbalanced, such that each picture served as a target and a foil equally often. Thus, if an infant tended to spend a lot of time looking at a given picture, irrespective of which picture had been labeled, it would not elevate their accuracy overall. Likewise, targets were presented on the left and right sides of the display equally often, reducing the likelihood that infants with tendency to look more to one side of the screen than the other might consistently but inadvertently look to the target picture due to its location rather than its correspondence with the label.

We suspected that one reason previous studies did not find relations between LPE and vocabulary size in infants under 15 months may be that LPE was assessed across words that posed a range of difficulty levels (e.g., Bergelson & Swingely, 2012; Zangl et al., 2005). If many of the words tested were unfamiliar to infants, the variability in performance on the trials testing them would reflect noise rather than true differences in comprehension, which would likely lead to poor sensitivity. We addressed this potential problem by separately assessing LPE on Easy and Hard words. We expected that performance on the Easy words would be best, and also that it would be most strongly related to LPE. Nonetheless, we also included Hard words so that we could evaluate both measures.

Parent Report Measure of Language Development. Parents filled out the MacArthur Communicative Development Inventory (MCDI), a commonly used assessment tool that provides information about infants' communicative development. We used the Words and Gestures version with 12-month-olds, focusing on measures of lexical development. Of primary interest were the measures of receptive and productive vocabulary size. These were assessed using a checklist on which parents indicate which words their infant understands or both understands and says. The MCDI can be used to generate raw counts of the number of words an infant is reported to understand and say. The raw counts can also be used to derive a percentile score, which reflects where infants' scores fall within a normed distribution for their age and sex. We have reported both measures in our analyses.

We expected that at 12 months, the MCDI receptive vocabulary size measures would provide the most useful measure of variability, given that most infants in our sample were producing very few words, if any, at this age (see Table 1). However, we also examined relations between LPE and productive vocabulary size at Visit 1. Most studies examining relations between vocabulary size and LPE in older infants have used productive vocabulary size, and the Words and Sentences MCDI form used with older infants in this study only includes a measure of productive vocabulary size. By including both receptive and productive vocabulary size at Visit 1 we were able to provide a fuller characterization of how LPE is (or is not) related to both vocabulary size measures over time.

Table 1: Vocabulary Size at Visits 1 and 2

	M	*SE*
Visit 1 Vocabulary Size		
MCDI Comprehension Raw	79.79	7.66
MCDI Comprehension %	43.27	3.71
MCDI Production Raw	8.03	1.35
MCDI Production %	52.19	2.49
Visit 2 Vocabulary Size		
MCDI Production Raw	68.98	16.34
MCDI Production %	41.35	4.74

Note: This table depicts information about infants' MCDI vocabulary size measures at each visit. Raw scores reflect the total number of words a parent reported their child to understand (Comprehension) or say (Production), and the "%" score reflect the percentile scores that corresponded to the raw scores.

Visit 2 Measures

LPE Task. Lexical recognition was tested using the materials and methods developed by Fernald and colleagues for use with infants 15 months and older. The target words we analyzed were "birdie", "baby", "car", and "shoe". The words "doggy" and "kitty" were also tested, but we excluded these items from the analyses. This allowed us to assess LPE using completely non-overlapping sets of words at the two visits, and thus any relations observed between the measures would not reflect familiarity with a specific set of words. The trials were structured very similarly to those in the Visit 1 LPE task. On each trial, an image appeared in the lower right and left corners of the screen and infants were asked to "Find the [target word]" after a 3 second silent baseline. Each of the words served as the target 4 times. All images were presented equally often, and side of presentation, both when serving as Targets and foils, was counterbalanced.

Parent Report Measure of Language Development. Parents filled out a MCDI questionnaire at Visit 2. The Words and Gestures version was used for 15-16 month-olds, and the Words and Sentences version was used for 17-19-month olds. The latter version only assesses productive vocabulary, with parents checking words that their child says. Furthermore, in studies with infants at this age it is common to test for relations between LPE and lexical development using a measure of productive vocabulary size. Thus we used only the production measure for all the infants at Visit 2. In addition, to reduce variability in this measure due to differences in infants' age, we focused on the percentile scores in our analyses, though we also report analyses using the raw scores to facilitate comparing our results to those from prior studies in which raw scores were used.

Results and Discussion

LPE Data Coding

Infant's eye movements during both the LPE tasks (at Visit 1 and Visit 2) were digitally recorded at a rate of 30 frames-per-second. Trained observers who were blind to the content of each trial coded the videos using the custom software iCoder. On each frame, coders indicated whether an infant was looking to the picture on the right, left, transitioning between pictures, or not attending to the display. Data from a quarter of the participants was randomly selected and recoded to assess reliability using a comparison function built into iCoder. The resulting agreement was 98% across all coded frames.

Visit 1 LPE

At Visit 1, we assessed infants' LPE in terms of both Accuracy and reaction time (RT). Accuracy is the typical measure used with infants younger than 15-18 months, reflecting infants' relative preference for the target picture after hearing it labeled. We assessed Accuracy across two time windows that were each 1500ms in duration: The Early window began 300ms after the onset of the spoken label, and ended at 1800ms, and the Late window captured the following 1500ms. The time frame captured by the Early Window is standardly used in studies with infants 15-months-old and older. However, because some studies with younger infants have assessed recognition over longer windows (e.g., Bergelson & Swingley, 2012), we also assessed Accuracy in the Late window. For each window (Early and Late) the Accuracy measure was computed by summing the number of frames spent looking to the target during a given trial, and dividing that number by the summed frames spent looking to the target and the distractor, which yielded a proportion score. This proportion score was averaged across trials of each type (e.g., Easy and Hard) to create an Accuracy score that reflected how much infants looked at the labeled picture in the Early and Late windows. Trials during which infants were not attending to the task for half of the silent baseline or half of the post-label windows were excluded. Infants had to contribute a minimum of 2 trials of a particular type for their data to be included (i.e., an infant with 5 Easy trials and 1 Hard trial would just contribute data for the Easy Accuracy measure). On average infants contributed 12.33 Easy trials (range 3-16), and 12.12 Hard trials (range 5-16). This yielded a total of 814

Easy Trials and 814 Hard trials for inclusion in the analyses. All 66 infants contributed data to the Easy Accuracy measure, and all but one contributed data to the Hard Accuracy measure.

We first tested whether infants showed evidence of recognizing Easy and Hard words. Using one-sample *t*-tests comparing Accuracy scores on each kind of trial to chance, or 0.5, we found evidence of comprehension for the Easy words in both the Early and Late windows (see Table 2 for the results of these *t*-tests; see also Figure 1). There was no evidence that infants recognized the Hard words in either window (Table 2; Figure 1). However, paired-sample *t*-tests revealed that performance on Easy trials was only marginally better than performance on Hard trials (Early Window t (64) = 1.590, p > .1; Late Window t (64) = 1.94, p < .1).

Table 2: LPE Performance at Visit 1

	N	***M***	***SE***	***t (df)***
Visit 1 LPE				
Early Window Easy Accuracy	66	0.538	0.015	2.56 (65)*
Late Window Easy Accuracy	66	0.541	0.015	2.63 (65)*
Early Window Hard Accuracy	65	0.508	0.015	0.576 (64)
Late Window Hard Accuracy	65	0.504	0.014	0.316 (64)

Note: * p < .05, ** p < .01. ***, p < .001. This table contains infants' mean performance on the LPE task, both in terms of their Accuracy and RT. The far right column depicts the results of one-sample two-tailed *t*-tests comparing Accuracy scores to chance (0.5).

Is Visit 1 LPE Related to Concurrent Vocabulary Size?

In the next set of analyses we tested whether Visit 1 LPE was related to MCDI measures of vocabulary size (see Table 3 for the relevant correlation coefficients). Infants' Accuracy in the Early window of Easy trials was related to the MCDI percentile scores for receptive vocabulary size (see Figure 2), and marginally to the raw receptive vocabulary size scores. No other relations between the Accuracy and vocabulary size measures approached significance. These findings suggest that infants who were more Accurate in the early portion of Easy trials also tended to understand more words.

To illustrate real-time recognition performance on Easy trials as a function of receptive vocabulary size, we divided infants into those with Small and Large receptive vocabularies using a median split on the normed scores. We then plotted their looking behavior during the Early and Late windows. These data, depicted in Figure 3, suggest that infants with relatively large receptive vocabularies performed better in the Early window on Easy trials than infants with smaller vocabularies. An independent-samples *t*-test comparing the two groups' performance in the Early window confirmed that infants with Large vocabularies were more accurate (M = .583, SE = .020) than infants with Small vocabularies (M = .494, SE = .020, t (64) = 3.20, p < .01). Furthermore, one sample *t*-tests comparing the Accuracy scores to chance (0.5) in the Early window revealed that infants with relatively large vocabularies showed evidence

Table 3: Relations Between Visit 1 Vocabulary Size and LPE Measures

	1	2	3	4	5	6	7
1. V1 MCDI Comprehension (raw)							
2. V1 MCDI Comprehension (%)	0.958***						
3. V1 MCDI Production (raw)	0.611***	0.552***					
4. V1 MCDI Production (%)	0.628***	0.619***	0.843***				
5.V1 LPE: Early Easy Accuracy	0.231Ŧ	0.306*	0.135	0.141			
6. V1 LPE: Late Easy Accuracy	0.024	0.023	0.128	0.024	0.333**		
7. V1 LPE: Early Hard Accuracy	-0.030	-0.013	0.107	0.105	0.157	.058	
8. V1 LPE: Late Hard Accuracy	-0.015	-0.059	0.074	-0.078	-0.013	0.178	0.320*

Ŧ p <.1, * p < .05, ** p < .01, *** p < .001

Note: This table contains the Pearson correlation coefficients for relations between infants' LPE (both Accuracy and RT) and MCDI Vocabulary size measures at Visit 2. The MCDI Raw scores reflect the total number of words a parent reported their child to understand or say, and the MCDI % scores reflect the percentile scores that corresponded to the raw scores.

of recognizing Easy words, but those with small vocabularies did not (see Table 4). While infants with larger receptive vocabularies showed evidence of successful recognition during the Late Window on Early trials, their performance was not better than that of infants with Small vocabularies in that window (t (64) = .338, p < .1.).

Recall that across the entire sample, infants did not show evidence of comprehending the Hard words. Furthermore, there was no evidence that performance on Hard word trials was correlated with vocabulary size. Consistent with these findings, performance on Hard trials was poor for both infants in the Small and Large vocabulary groups (see Figure 4, and Table 4), with no evidence that Accuracy scores differed from chance (0.5) for either group in the Early or Late windows (Table 4). The two groups' Accuracy scores also did not differ from each other (Early Window t (63) = .378, p > .1; and Late Window t (63) = -.553, p > .1).

Table 4: LPE Performance at Visit 1 in Infants with Large and Small Vocabularies

	N	***M***	***SE***	***t (df)***
Large Vocabulary				
Early Window Easy Accuracy	33	0.583	0.020	4.253 (32)***
Late Window Easy Accuracy	33	0.547	0.019	2.467 (32)*
Early Window Hard Accuracy	33	0.514	0.021	0.627(32)
Late Window Hard Accuracy	33	0.497	0.021	-0.154 (32)
Small Vocabulary				
Early Window Easy Accuracy	33	0.494	0.020	-.0323(32)
Late Window Easy Accuracy	33	0.536	0.026	1.414(32)
Early Window Hard Accuracy	32	0.502	0.011	0.130 (31)
Late Window Hard Accuracy	32	0.512	0.018	0.662 (31)

Ŧ p <.1, * p < .05, ** p < .01, *** p < .001

Note: This table contains mean performance on the LPE task for infants in the Large and Small Vocabulary groups. The far right column depicts the results of one-sample two-tailed *t*-tests comparing Accuracy scores to chance (0.5).

Altogether, these findings suggest that 12-month-olds successfully recognized the commonly-known words (i.e., words reported to be understood at this age by more than 50% of parents) we tested them, but did not show reliable evidence of recognizing the Hard words, which are estimated to be understood by only about a quarter of 12-month-olds. However, this effect appeared to be carried by infants with larger vocabularies, as they showed evidence of recognizing the Easy words, while the infants with relatively small vocabularies did not. Correlational analyses also revealed infants who are better able to recognize these relatively early-learned words have larger receptive vocabularies according to the MCDI. Interestingly, even though infants with larger vocabularies showed above-chance looking to the labeled pictures during both the Early and Late widows, only differences in their Accuracy during the Early Window were related to differences in vocabulary size.

While these results suggest there is a relation between LPE and vocabulary size, by performing multiple comparisons we inflated the possibility of a Type 1 error. Thus, we note that it is important to view the larger pattern of results before strongly interpreting these results as evidence for an early emerging relationship between LPE and vocabulary size. In particular, if infants' Accuracy on Easy trials is also related to LPE measures and Vocabulary size at Visit 2, we would have stronger evidence to suggest that these individual differences in performance on the LPE task are related to lexical development.

Visit 1 RT

RT is not often used to assess LPE in infants under 15 months. Indeed, Fernald and colleagues (2008) advise against it because RT tends to be quite noisy in younger infants. However, we included RT measures for both Easy and Hard trials to determine whether it is unreliable using the current materials. Because we found infants with relatively large and small vocabularies differed in their performance on the LPE task during the Early window, it is likely that including shifts within this window in the RT calculation would provide the most sensitivity. Thus, RT was calculated as an average of the time taken for infants to initiate a shift from the distractor picture to the target if that shift occurred between 300 to 1800 ms after the label onset. Using this cutoff, which is standard for older infants, also allowed us to compare RT at Visit 1 and Visit 2 when computed using the same window. Note that many fewer infants contributed sufficient RT trials to be included in these analyses: 45 infants contributed RT data for Easy trials and 36 contributed RT for Hard trials. Those with usable data contributed between 2 and 6 trials (M = 2.76; for a total of 152 trials included in the analyses) and between 2 and 8 Hard trials (M = 2.82, and a total of 151 trials). Infants' mean RT on Easy trials was 946.85 ms (*SE* = 38.09) and mean RT on Hard trials was 957.92 ms (*SE* = 43.41). RT on Easy and Hard trials did not differ (t (30) = .272, $p > .1$).

RT is computed over a smaller number of trials than the Accuracy measure, and captures time to shift to the target picture rather than duration of looking to it. Nonetheless, infants who are faster to shift to the target picture (i.e., those with faster RTs) are also likely to have higher accuracy scores (i.e. they are also more likely to be looking at the target for longer portions of the target windows). Indeed, RT on Easy trials was correlated with Accuracy in the Early window of Easy trials (r (44) = -.551, $p < .001$) and RT on Hard trials was correlated with Hard Accuracy in the Early window (r (35) = -.358, $p < .05$) though less strongly. In contrast to the Accuracy measure, however, neither RT on the Easy nor the Hard trials was significantly related to any of the vocabulary size measures (Table 5). We note, though, that RT on Easy trials was marginally correlated with normed receptive and productive vocabulary size. Thus, both RT and Accuracy measures of LPE for commonly known words during the Early window showed signs of being related to receptive vocabulary size at 12 months of age, but the effects were only significant for the Accuracy measure.

Table 5: Relations Between Reaction Time and Vocabulary Size at Visit 1

	RT Easy	RT Hard
V1 MCDI Comprehension (raw)	-0.166	-0.037
V1 MCDI Comprehension (%)	-0.290Ŧ	-0.072
V1 MCDI Production (raw)	-0.117	-0.239
V1 MCDI Production (%)	-0.255Ŧ	-0.272

Ŧ p <.1, * p < .05, ** p < .01, *** p < .001

Note: This table contains Pearson correlation coefficients reflecting the concurrent associations between RT of the LPE task and vocabulary size at Visit 1. The MCDI Raw scores reflect the total number of words a parent reported their child to understand or say, and the % scores reflect the percentile scores that corresponded to the raw scores.

What Measures of LPE are related in Vocabulary Size at Visit 1 in Infants Contributing Visit 2 Data?

Before testing whether Visit 1 LPE measures were related to Visit 2 LPE and vocabulary size, we tested whether the key relations we observed between LPE and vocabulary size measures held in the sample of infants who contributed usable data LPE at Visit 2. Our aim was to determine whether the measures that were best predictors of concurrent vocabulary size in the full sample were also the best predictors in the subsample of infants who contributed data at Visit 2.

When we compared the MCDI and LPE scores of infants who did and did not contribute usable data at Visit 2, we did not find any significant differences between the groups. However, one difference was marginally significant: Infants who contributed data at Visit 2 tended to have higher Accuracy scores on Easy words during the Early Window (*M* = 0.561, *SE* = 0.019) than infants who did not

(M = 0.503, SE = 0.023); t (64) = 1.921, p = .059. The range of scores, however, was comparable, with Accuracy scores falling between 0.3 and 0.77 in infants who did not contribute Visit 2 data, and between 0.36 and 0.88 in infants who did.

In the sample of infants who contributed LPE data at Visit 2, Visit 1 Easy Accuracy in the Early window was related to the percentile score for MCDI receptive vocabulary size (r (39) = 0.343, p < .05), just as it was in the full sample. RT on the Easy trials was significantly correlated with receptive percentile scores (r (26) = -0.401, p < .05), and with both raw and percentile productive vocabulary size (rs (26) > -0.394, ps < .05). There were no other differences in the patterning of the relations between Accuracy, RT, and MCDI scores when considering the whole sample and the smaller sample of infants contributing Visit 2 data.

In sum, Accuracy on Easy trials was the measure that most strongly correlated with vocabulary size in the full sample, and these measures were also correlated in infants who contributed Visit 2 data. RT on Easy trials, while only marginally related to MCDI scores in the full sample, was significantly correlated with vocabulary size in infants who contributed Visit 2 data, who also had relatively high Accuracy scores.

Visit 2 LPE

At Visit 2 LPE was assessed in terms of Accuracy and RT using looking behavior only during the Early window (the window including 300 to 1800ms after the onset of the target label). This is the standard window for computing these measures in infants aged 15 months and older. We used looking behavior in this window to compute both Accuracy and RT measures using the same parameters and calculations as above. Infants contributed from 7 to 24 usable trials (*M* = 19.31) to the Accuracy measure, yielding a total of 753 trials for inclusion across the 40 infants contributing data. Summing across the 34 infants contributing sufficient RT data for inclusion yielded a total of 158 trials for use in the RT analysis, with infants contributing anywhere between 2 and 15 usable trials (*M* = 4.65). Note that there were 6 infants who contributed sufficient usable trials to compute an Accuracy score, but who did not have a usable RT score. Their data are included in all of the analyses except those using the Visit 2 RT measure.

Infants showed evidence of recognizing the tested items, with Accuracy scores well above chance, or 0.50 (M = 0.601, SE = .020; t (39) = 4.979, $p < .001$; see also Figure 1), as we expected they would based on prior work using this task with infants in this age range (e.g., Fernald et al., 2006). Their mean RT was 910.82 ms (SE = 45.70). RT and Accuracy were strongly correlated (r (33) = -0.645, $p < .001$) at this age.

<u>Is Visit 2 LPE Related to Concurrent Vocabulary Size?</u>

We replicated previous findings that RT on the LPE task is related to concurrent vocabulary size, as assessed by the MCDI, at Visit 2 (see Table 6). RT was most strongly related to raw productive vocabulary size scores, but it was also marginally correlated with the normed productive vocabulary scores. However, Visit 2 Accuracy scores were not related to either measure of productive vocabulary size. It is more common to use RT as a measure of LPE than Accuracy in infants at this age, and it was the LPE measure most strongly associated with concurrent vocabulary size in our sample.

Table 6: Concurrent Relations Between LPE and Vocabulary Size at Visit 2

	1	2	3	4
1. V2 Age				
2. V2 LPE Accuracy	0.270 Ŧ			
3. V2 LPE RT	-0.252	-0.645***		
4. V2 MCDI Production Raw	0.291Ŧ	0.150	-0.373*	
5. V2 MCDI Production %	-0.032	0.012	-.0299Ŧ	0.779***

Ŧ p <.1, * p < .05, ** p < .01, *** p < .001

Note: This table depicts the Pearson correlation coefficients for relations between key Visit 2 (V2) measures. The MCDI Raw scores reflect the number of words that parents report them to say, and the Production % reflects the age-normed score. There were 40 infants contributing scores to all analyses except those involving RT, where there were 34.

Is there Stability in Vocabulary Size and LPE Measures Over Time?

We next asked whether there was continuity in infant’s’ vocabulary size and LPE across the visits. An infants’ MCDI scores are typically related over time, such that earlier and later administrations within and across the Words and Gestures and Words and Sentences forms typically correlate at fairly high levels (Fenson et al., 1994; 2007). As can be seen in Table 7, this was the case in our sample as well. Infants' MCDI scores at Visit 2 were comprehensively related to those at Visit 1, with receptive and productive measures correlating with themselves and with each other over time.

Table 7: Relations Between Vocabulary Size Across Visits 1 and 2

	V2 MCDI Production (raw)	V2 MCDI Production (%)
V1 MCDI Comprehension (raw)	0.45**	0.52***
V1 MCDI Comprehension (%)	0.53***	0.58***
V1 MCDI Production (raw)	0.48**	0.50***
V1 MCDI Production (%)	0.45**	0.51**

** $p < .01$, *** $p < .001$

Note: This table contains Pearson correlation coefficients reflecting the associations between MCDI measures across visits. The MCDI Raw scores reflect the number of words that parents report them to comprehend or produce, and the % scores reflect the corresponding normed scores.

We were particularly interested in whether there was continuity in infants' performance on the LPE tasks across visits, as this could help us evaluate whether the individual differences we observed in Accuracy and RT at 12 months are meaningful. In particular, if infants who have better LPE at 12 months also tend to have better LPE at Visit 2, the latter reflecting an age range in which there is strong evidence that individual differences in LPE are meaningful, it would suggest that these differences are already emerging at 12 months.

If individual differences in LPE at 12 months are stable, one possible result would be that a given measure of LPE would be related across visits over time. For example, Visit 1 Accuracy on Easy trials might be correlated with Visit 2 Accuracy. However, if Accuracy on the Easy trials is a more sensitive measure than RT in younger infants, and RT is more sensitive in older infants, it is possible that these measures are most strongly correlated across visits. As can be seen in Table 8 and Figure 5, this was the case: Infants who were more accurate in the Early Window of Easy trials at Visit 1 had faster RT scores on the LPE task at Visit 2. Faster RT on Hard trials at Visit 1 also marginally predicted faster RT at Visit 2. However, no other Visit 1 measures, including RT on the Easy trials, predicted either of the Visit 2 LPE measures. Note that these analyses only include infants who contributed data at Visit 2, and for whom both Accuracy and RT were related to concurrent vocabulary size. These findings suggest that at Visit 1, infants' Accuracy on Easy trials during the Early Window may be a more robust measure of individual differences in LPE than RT.

However, our data also suggest that RT becomes a better measure of LPE than Accuracy as infants approach 18 months: Visit 2 RT was correlated with both Visit 1 Accuracy and also Visit 2 vocabulary size, while Visit 2 Accuracy was not related to any of the other measures, suggesting it is the more sensitive measure of LPE at Visit 2. In fact, many studies use RT over Accuracy as a measure of LPE at this age.

Table 8: Relations Between LPE Measures Across Visits 1 and 2

Visit 1 LPE	Visit 2 LPE Accuracy	Visit 2 LPE RT

Early Window Easy Accuracy	0.041	-0.373*
Late Window Easy Accuracy	0.158	-0.194
Early Window Hard Accuracy	0.154	-0.157
Late Window Hard Accuracy	-0.021	0.074
Easy RT	-0.047	0.302
Hard RT	-0.328	0.353Ŧ

Ŧ p <.1, * p < .05, ** p < .01, *** p < .001

Note: This table contains Pearson correlation coefficients reflecting the association between LPE and MCDI measures across visits.

Does Visit 1 LPE Predict Visit 2 Vocabulary Size?

A key question we wanted to address was whether individual differences in LPE at 12 months predict later vocabulary size. To that end, we tested whether measures of Accuracy and RT on the LPE task at 12 months predicted MCDI estimates of productive vocabulary size on the MCDI at Visit 2 (see Table 9 for relevant correlations). Here we only included Accuracy in the Early Window (not the Late Window) as a measure of Visit 1 LPE, given evidence that the individual differences in this measure were concurrently related to vocabulary size, and also predicted Visit 2 RT. Because infants' age and raw vocabulary size were marginally correlated at Visit 2 (see Table 6), we focused on predicting MCDI percentile scores. Recall that the percentile scores are derived by assigning a score that reflects where their raw score falls in the distributional of scores for infants at that age and sex, which promotes comparison across ages. In fact, age was unrelated to the normed scores (see Table 6). Thus, using the percentile scores allows us to reduce the influence of variance in infants' age at Visit 2 on the measure of their vocabulary size. We also computed a measure of change in the MCDI productive vocabulary size percentile scores by subtracting infants' Visit 1 productive vocabulary size percentile scores from those at Visit 2. This allowed us to assess whether infants who had begun to produce relatively more words for their age and sex across the visits had better LPE at Visit 1.

The results of Pearson correlations revealed that higher Accuracy scores on Easy trials at Visit 1 predicted larger productive vocabularies at Visit 2 (Table 9). Infants with higher Accuracy scores on Easy trials were also more likely to experience increases in productive vocabulary size percentile scores between Visit 1 and Visit 2 (see Figure 6). Infants' Visit 1 Accuracy on the Hard Trials, did not predict any Visit 2 MCDI measures, nor did it predict changes in those measures.

Recall that Visit 1 RT on Easy trials was marginally correlated with concurrent receptive and productive vocabulary size in the sample as a whole, and was significantly related to these measures of vocabulary size in infants who contributed Visit 2 data. In the current analyses we found that Visit 1 RT on Easy trials was also correlated with Visit 2 productive vocabulary size percentile scores, but not with the change score (Table 9).

To summarize, there was substantial evidence that the Visit 1 Accuracy on Easy trials captured individual differences in LPE: It was correlated with concurrent vocabulary size percentile scores, both in the sample as a whole and in the subsample that contributed LPE data at Visit 2, and it was also correlated with LPE and vocabulary size percentile scores several months later.

In contrast, RT on Easy trials at Visit 1 was related to concurrent vocabulary size, but only in the subset of the infants who contributed Visit 2 data, who tended to have higher Accuracy scores. RT on Easy trials at Visit 1 was correlated with vocabulary size at Visit 2, but was not related to Visit 2 LPE. While the findings with RT are suggestive, and converge with the findings with Accuracy to suggest that individual differences in lexical recognition at 12 months are meaningful we urge caution in interpreting the results to suggest that RT should be used to measure of LPE at 12 months, as Visit 1 Accuracy was generally more strongly and consistently related to measures of vocabulary size and later LPE than Visit 1 RT.

Table 9: Relations Between Visit 1 LPE and Visit 2 Vocabulary Size

	V2 MCDI Production Raw	V2 MCDI Production %	V1-V2 Change in Production %
V1 LPE: Easy Accuracy	0.266Ŧ	0.361*	0.321*
V1 LPE: Hard Accuracy	0.125	0.247	0.259
V1 LPE: Easy RT	-0.300	-0.423*	-0.215
V1 LPE: Hard RT	-0.378 Ŧ	-0.179	0.042

Ŧ p <.1, * p < .05, ** p < .01, *** p < .001
Note: This table contains Pearson correlation coefficients reflecting the association between Visit 1 LPE measures and Visit 2 MCDI measures across visits. The V1-V2 Change score was computed by subtracting infants' productive vocabulary percentile score at Visit 1 from their productive vocabulary percentile score at Visit 2, with positive values indicating increases in vocabulary size.

Does Visit 1 Vocabulary Size Predict Visit 2 LPE?

The primary goal of the current study was to test whether relations that hold between LPE and vocabulary size at 18 months, once lexical development is well underway, also hold in younger infants. We were particularly interested in testing whether LPE is related to concurrent and later vocabulary size at 12 months because the existence of these relations in older infants is well-documented, and have been taken as evidence that LPE contributes to gains in vocabulary size (Fernald & Marchman, 2012). However, as described earlier, the relation between LPE and vocabulary size is likely to be bidirectional. Some of the clearest evidence for the possibility that amassing a relatively large vocabulary may promote LPE has come from a study using a substantially different design (e.g., Borovsky et al., 2016). Nonetheless, in the current work we can test whether having a large vocabulary is related to later LPE. To that end, we tested whether vocabulary size at Visit 1 predicts measures of LPE at Visit 2, as well as change in them from Visit 1 to Visit 2. We used LPE Easy Accuracy in the Early window as the Visit 1 Accuracy measure in that calculation because it appeared to be most sensitive to individual differences in LPE at that age. However, we found that infants' vocabulary size at Visit 1, no matter how it was assessed, was not significantly related to LPE at Visit 2 (see Table 10). However, there was one marginal correlation consistent with the possibility that having a larger vocabulary is related to later LPE. Specifically, infants with larger raw productive vocabularies at Visit 1 tended to have better Visit 2 LPE Accuracy scores. This weak relation may be meaningful, though it should be interpreted with caution given that the specific measures involved did not appear to be the more sensitive measures of individual variability in either vocabulary size or LPE in our sample.

Table 10: Relations Between Visit 1 Vocabulary Size and Visit 2 LPE

	V2 LPE Accuracy	V2 LPE RT	Change in LPE Accuracy from V1 to V2	Change in LPE RT from V1 to V2
V1 MCDI Comprehension (raw)	-0.09	-0.03	-0.19	0.24
V1 MCDI Comprehension (%)	-0.04	-0.11	-0.23	0.19
V1 MCDI Production (raw)	0.28Ŧ	-0.13	0.17	0.10
V1 MCDI Production (%)	0.14	-0.07	0.07	0.21

Ŧ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Note: This table contains Pearson correlation coefficients reflecting the associations between Visit 1 vocabulary size measures and Visit 2 LPE measures. The measure of V1 Accuracy used in these analyses was based on performance in the Early Window on Easy Trials. The Accuracy Change score was computed by subtracting Visit 1 Accuracy scores on those trials from Accuracy at Visit 2, and the RT Change score was computed by subtracting Visit 1 RT on Easy trials from Visit 2 RT.

General Discussion

By the time they are 18 months old, infants' lexical-processing efficiency (LPE) is related to how many words they are saying, and predicts which late-talking infants are likely to catch up to their typically-developing peers (Fernald & Marchman, 2012). There is also evidence that growth in the lexicon supports LPE, as infants are faster to recognize words that come from denser semantic networks (Borovsky et al., 2016). These findings suggest that a synergy between lexical knowledge and real-time lexical processing supports growth in the lexicon across the following years.

However, in prior work with 6- and 14-month-olds, individual differences in performance on lexical recognition tasks were not reliably correlated with measures of vocabulary size (e.g., Bergelson & Swingley, 2012; Reznick, 1990; Zangl et al., 2005). These null effects raised questions about when individual differences in LPE emerge, and whether real-time processing skill supports vocabulary growth early in lexical development. Characterizing the development of LPE and its relation to vocabulary size is relevant to determining whether the striking gains in lexical knowledge and processing skill in the latter part of the second year reflect the operation of a very different mechanism that supports early lexical recognition and word learning processes.

To that end, in the current work we tested whether LPE is related to vocabulary size in 12-month-olds when we separately assessed LPE using words that infants are more vs. less likely to know according to parental report. We suspected that measures of LPE based on relatively familiar words would yield a measure with greater sensitivity. Another important dimension of our design involved testing infants' LPE and vocabulary size a few months later when they were at least 15 months old. This allowed us to assess whether there is stability in LPE measures within infants over time, and also whether measures of LPE at 12 months are related to concurrent and later vocabulary size.

Across several analyses we found evidence that individual differences in LPE are present by 12 months of age. First, 12 months olds who showed better real-time comprehension of commonly known words (i.e., those who had higher Accuracy, spending more time looking to the target pictures in the period just after they were labeled) tended to have larger receptive vocabularies as assessed by the MCDI. The relation was modest, only holding between Accuracy on words that a majority of 12-month-olds are likely to know, and their receptive vocabulary size percentile scores. Critically, this Accuracy measure of LPE at 12 months was also correlated with how quickly infants found the referents of a different set of early-learned words several months later at Visit 2 (i.e., with Visit 2 RT). Moreover, 12-month-olds' Accuracy on these trials predicted their productive vocabulary size percentile scores at Visit 2. Altogether these results suggest that individual differences in 12 month-olds' LPE are meaningful and stable. Furthermore, they suggest that the relation between LPE and vocabulary size is established fairly early in lexical development.

Measuring Individual Differences in LPE

These data may help researchers interested in assessing younger infants' LPE in choosing a good measure. Specifically, because we included multiple measures of LPE (i.e., Accuracy and RT), we were able to evaluate which of them showed evidence of relating to concurrent and later vocabulary size, and which related to LPE several months later. The measure that was most consistently related to current and later vocabulary size was infants' Accuracy on Easy trials: As stated above, it was concurrently related to receptive vocabulary size percentile scores, and predicted productive vocabulary size percentile scores, as well as increases in them, several months later. Accuracy on Easy trials also predicted RT at Visit 2, which is a commonly-used measure of LPE at 15-18 months. Infants' Accuracy on Hard trials was not related to any MCDI or later LPE measures, suggesting that using words that are likely to be familiar to most infants yields an LPE measure that is more sensitive to individual differences. The Appendix reports analyses using a related measure of target looking, computed by taking advantage of the yoked structure of the trials. The results of analyses using this measure did not yield substantively different results than those using Accuracy scores.

Infants' RT on Easy trials was strongly correlated with Accuracy on Easy trials, however the evidence that this measure captured variability in lexical recognition skill was somewhat weaker. For example, infants' RT on Easy trials was only marginally related to concurrent vocabulary size. It was significantly related to Visit 2 productive vocabulary size, but it was not related to Visit 2 RT. Infants' RT on Hard trials was marginally related to RT at Visit 2, but to nothing else. Interestingly, Visit 1 RT on Easy trials was significantly related to concurrent receptive and productive vocabulary size in infants who returned for Visit 2. These infants tended to have higher accuracy scores, suggesting that RT at 12 months may be a less sensitive measure in samples including infants with a wider range of language skills, especially infants at risk for language learning difficulties.

Overall our data suggest that the RT measure may be a less sensitive and/or reliable measure of individual differences than Accuracy at this age. This is likely because it was computed over substantially fewer trials, and fewer infants contributed sufficient usable trials to be included in the analyses. It might be possible to increase the sensitivity of RT in younger infants by including more trials, though an advantage to using Accuracy is that it allows for more efficient assessment of LPE.

Limitations

While these data suggest that there is continuity in lexical recognition skill, and that it may support vocabulary development as infants approach 18 months, there are some caveats to these conclusions. One is that our sample was relatively homogenous. Infants were primarily Caucasian, and were growing up in families with high levels of maternal education. Thus, we cannot generalize our findings beyond this sample, and it will be important to test these questions in a more heterogeneous group of infants. Nevertheless, infants exhibited a range of vocabulary scores on a normed measure (the MCDI), with mean scores below the 50th percentile. This suggests that our findings are not limited to infants who are relatively advanced in their language development for their age.

Another issue is that we did not obtain and usable follow-up data from more than 1/3 of our participants assessed at Visit 1, either because they did not return to the lab for a second assessment or because the data we collected from them when they returned could not be included because of fussiness, inattention, etc. Infants who contributed data at the follow-up visit tended to have better Accuracy on the LPE task, suggesting that the follow-up data over-represented infants with relatively good performance on the LPE task, again highlighting the need to replicate these findings in future work with larger and more diverse samples. Furthermore the age at follow-up was not tightly constrained. However, we were able to partially address variance in age by using a percentile score for our key outcome measure (vocabulary size) that normalizes across age.

Future Directions

Despite these limitations, our results suggest that by 12 months of age, performance on LPE tasks can capture variation in real-time language processing skills, and that infants who are better able to recognize familiar words as they unfold in real time are already at an advantage for learning the words in their language. One way to gain traction in understanding how LPE may support early lexical growth would be to test whether and how it relates to word-learning processes. At least one study has already begun to ask this question in older infants (Lany, 2017). For example, by 17 months of age, infants who have better lexical recognition skills are better able to learn novel words in a fairly straightforward task (i.e., when labels are spoken in common ostensive labeling contexts such as “Look, it’s a ____”) as a single object is presented). By 30 months LPE no longer predicts word learning in this simple task, but it does when they need to use distributional cues within the sentence to establish reference (Lany, 2017). These findings suggest that LPE may support word-learning when it is challenging, and perhaps especially when the words surrounding a novel label provide crucial cues to meaning.

LPE may facilitate learning novel words in similar ways earlier in development. For example, by 18 months infants are faster to recognize familiar words when they are presented in common ostensive labeling frames vs. when they are presented in isolation (Fernald & Hurtado, 2006). Likewise, infants with better LPE may be better able to capitalize on the presence of ostensive labeling frames in learning novel words. In addition, LPE may also be especially helpful to novice word learners by promoting accurate encoding and robust memory for word forms. This would to aid infants in learn similar sounding words, such as minimal pairs, and using cross-situational cues to aid in establishing reference. LPE may be *less* relevant however, when cues from outside of spoken language are most relevant, such as the referent of a gaze or deictic gesture, such as indexical pointing.

Another important task for future work will be to identify the underpinnings of LPE itself. Developing larger, more densely connected lexical networks is likely to support improvements in LPE (Borovsky et al., 2016). However, there are likely to be other important factors, especially prior to amassing a relatively large vocabulary. By the time infants are 18-months, there is an association between hearing more language at home and strong lexical processing skills. A recent study with younger infants, however, reported that language input and word-form recognition skills are not correlated, and may have independent effects on later vocabulary size (Newman, Rowe, & Ratner, 2016). Thus, future work will be important to clarify whether qualities of infants' language environment impact the development of LPE in the first year of life.

Another possibility is that infants' ability to learn predictive structure in their language input supports developments in LPE. Recent work suggests that infants' ability to use statistical regularities (e.g., transitional probabilities between syllables) to segment fluent speech is related to their LPE by 15 months (Lany et al., 2017). Given evidence that infants can use sequential statistics to segment speech by 6-8 months (Saffran, Aslin, & Newport, 1996), or even earlier, it may be that sensitivity sequential structure, by making speech more predictable supports early lexical recognition. However, we note that LPE is likely to draw on multiple cognitive underpinnings, including object perception and categorization processes, as well as memory processes necessary for forming word-referent linkages. Thus, it would also be informative to determine whether infants' developing visual recognition and categorization skills support LPE.

In conclusion, previous work linked real-time processing to word-learning trajectories, but left open the possibility that real-time skill supports word learning only after lexical development is well underway. Here we found that meaningful individual differences in real-time comprehension are present by 12 months, correlating with receptive vocabulary size even when infants know relatively few words. Critically, 12-month-olds' real-time processing also predicted how quickly they added new words to their productive lexicons over the following months. These data suggest that real-time processing skills may support rapid gains in vocabulary size across the second year.

References

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9), 3253–3258.

Bergelson, E., & Swingley, D. (2013). The acquisition of abstract words by young infants. *Cognition*, *127*(3), 391–397.

Bergelson, E., & Swingley, D. (2015). Early word comprehension in infants: Replication and extension. *Language Learning and Development*, *11*(4), 369-380.

Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016). Semantic structure in vocabulary knowledge interacts with lexical and sentence processing in infancy. *Child Development*, *87*(6), 1893–1908.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, *28*(1), 125–127.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., & Stiles, J. (1994). Variability in early communicative development. *Monographs of the society for research in child development*, i-185.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., & Reznick, S. J. (2007). MacArthur–Bates Communicative Development Inventories: User's Guide and Technical Manual, 2nd edn (Baltimore, MD: Paul Brookes).

Fernald, A., & Hurtado, N. (2006). Names in frames: Infants interpret words in sentence frames faster than words in isolation. *Developmental science*, *9*(3).

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*, *83*(1), 203–222.

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, *16*(2), 234–248.

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*(1), 98.

Fernald, A., Pinto, J. P., Swingley, D., Weinbergy, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, *9*(3), 228–231.

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language. *Developmental Psycholinguistics: On-Line Methods in Children's Language Processing*, 113–132.

Fisher, C., Gleitman, H., & Gleitman, L.R. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology, 23,* 331-392.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2016). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, 1–18.

Junge, C., Kooijman, V., Hagoort, P., & Cutler, A. (2012). Rapid recognition at 10 months as a predictor of language development. *Developmental science*, *15*(4), 463-473.

Jusczyk, P. W., & Aslin, R. N. (1995). Infants′ detection of the sound patterns of words in fluent speech. *Cognitive psychology*, *29*(1), 1-23.

Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*, 606-608.

Lany, J. (2017). Lexical-processing efficiency leverages novel word learning in infants and toddlers. *Developmental Science*.

Lany, J., Shoaib, A., Thompson, A., & Estes, K. G. (2017). Infant statistical-learning ability is related to real-time language processing. *Journal of Child Language*, 1-24.

Law, F., & Edwards, J. R. (2015). Effects of vocabulary size on online lexical processing by preschoolers. *Language Learning and Development*, *11*(4), 331–355.

Marslen-Wilson, W., & Zwisterlood, P. (1989). Accessing spoken words: The importance of word onsets. Journal of Experimental Psychology: Human Perception and Performance, 15, 576–585.

Newman, R. S., Rowe, M. L., & Ratner, N. B. (2016). 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*, *43*(5), 1158–1173.

Reznick, J. S. (1990). Visual preference as a test of infant word comprehension. *Applied Psycholinguistics*, *11*(2), 145–166.

Singh, L., Morgan, J. L., & White, K. S. (2004). Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language*, *51*(2), 173-189.

Singh, L., Reznick, S. J., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: a longitudinal analysis. *Developmental Science*, *15*(4), 482–495.

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*(5), 480–484.

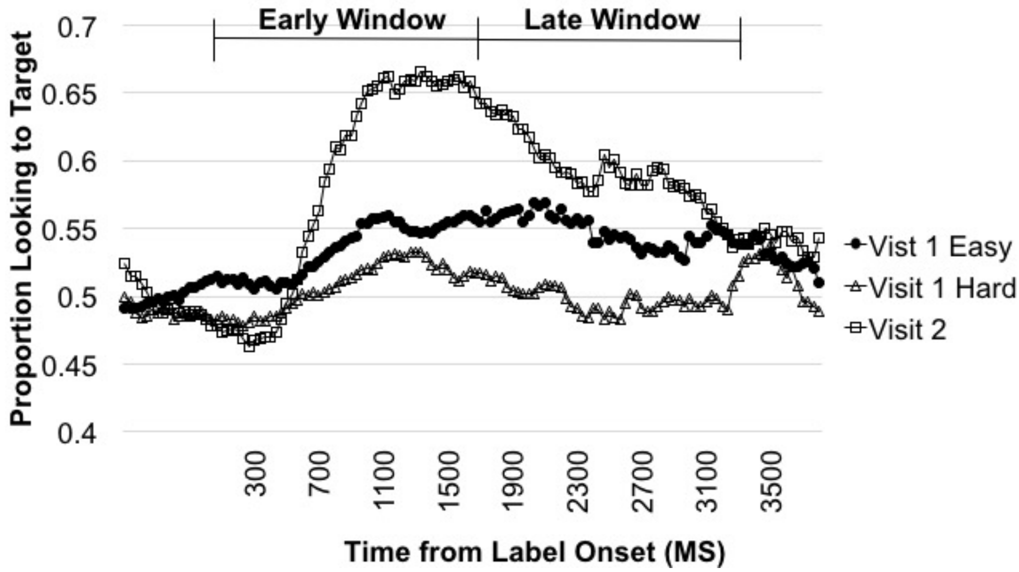
Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child development*, *75*(4), 1067-1084.

Weisleder, A., & Fernald, A. (2013). Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary. *Psychological Science*, *24*(11), 2143–2152. https://doi.org/10.1177/0956797613488145

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and development*, *7*(1), 49-63.

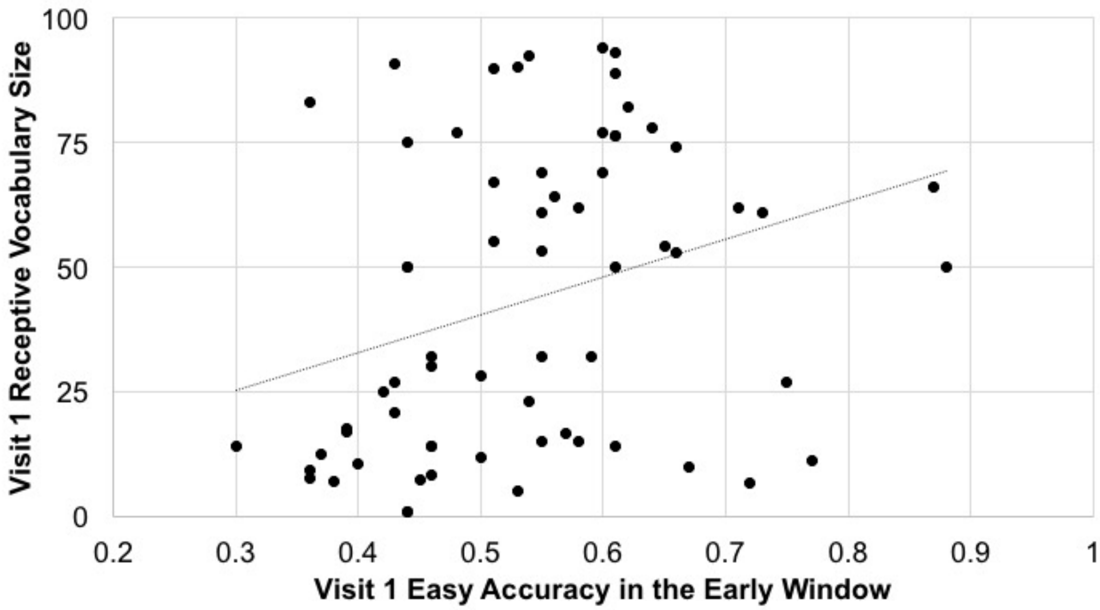
Zangl, R., Klarman, L., Thal, D., Fernald, A., & Bates, E. (2005). Dynamics of word comprehension in infancy: Developments in timing, accuracy, and resistance to acoustic degradation. *Journal of Cognition and Development*, *6*(2), 179–208.

Figure 1



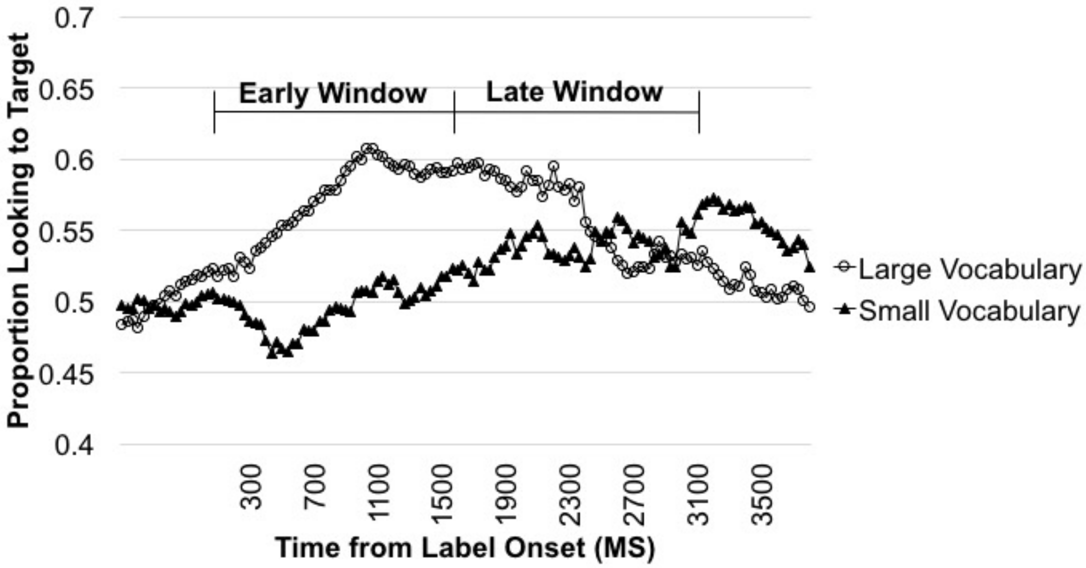
Caption: LPE at Visit 1 for Easy and Hard words. The label was played at 0ms and the analysis window was from 300 to 1800ms.

Figure 2



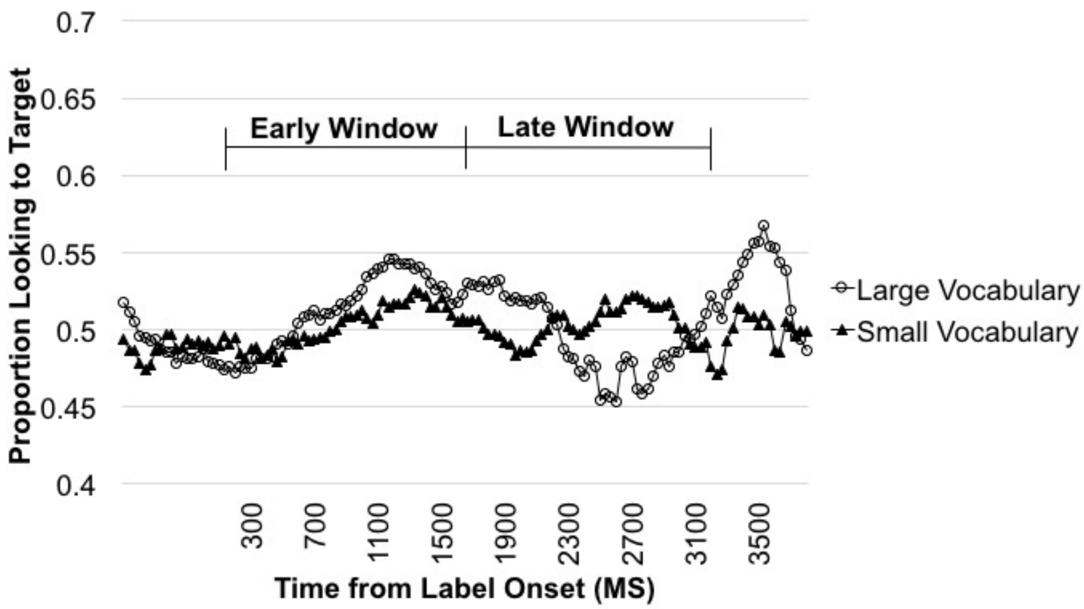
Caption: Infants with better Visit 1 LPE, as assessed by Accuracy in the Early Window on Easy trials, also had larger receptive vocabularies, according to the MCDI percentile scores.

Figure 3



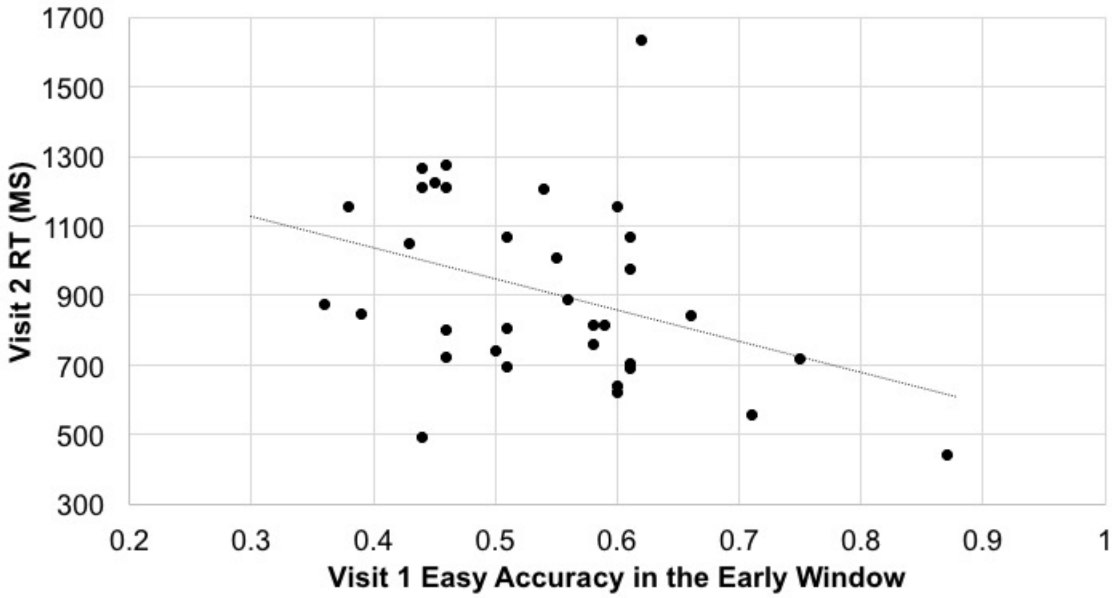
Caption: The figure depicts the proportion of trials infant were looking to the target picture, frame-by-frame, on Easy trials across the Early and Late Windows. Infants were grouped function of a median split on infants' concurrent receptive vocabulary size (as assessed by MCDI percentile scores), into Large and Small Vocabulary groups. The label for the target picture was played at 0ms and the analysis window was from 300 to 1800ms.

Figure 4



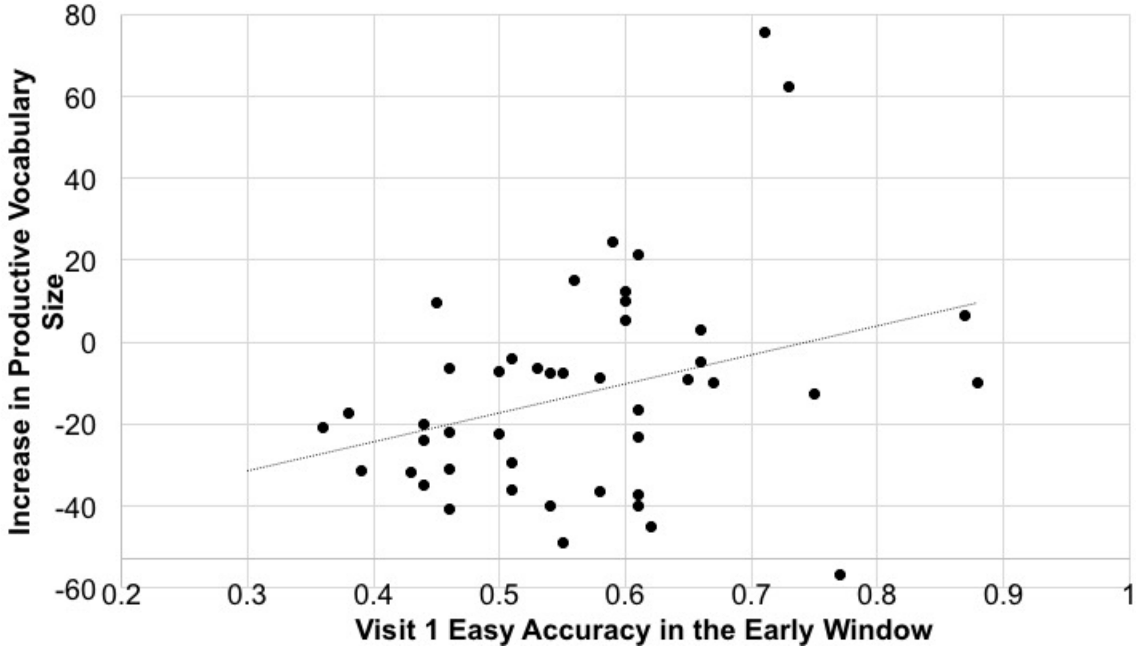
Caption: The figure depicts the proportion of trials infant were looking to the target picture, frame-by-frame, on Hard trials across the Early and Late Windows. Infants were grouped function of a median split on infants' concurrent receptive vocabulary size as measured by the MCDI percentile scores. The label for the target picture was played at 0ms and the analysis window was from 300 to 1800ms.

Figure 5



Caption: Infants with better Visit 1 LPE, as assessed by Accuracy in the Early Window on Easy Trials, were faster to recognize words in the LPE task at Visit 2.

Figure 6



Caption: Infants with better Visit 1 LPE, as assessed by Accuracy in the Early Window on Easy Trials, were experienced greater gains in productive vocabulary size percentile scores when assessed Visit 2.

Appendix: Supplementary Analyses

In the analyses reported in the Results section, we found that LPE at 12 months, when assessed using an accuracy measure that reflected the proportion of time infants spent looking to the target after the label was played, reflected meaningful individual differences in infants’ lexical development. Because the items on the LPE task were always presented in yoked pairs (e.g., "dog"-"milk"), we were also able to calculate a “Yoked" accuracy measure that reflected selective looking to the target when it was labeled. This yoked measure reflected the mean proportion of trials that infants were looking to one member of a picture pair, such as the dog, when it was labeled, after subtracting their mean proportion looking to dog when the other member of the pair (i.e., "milk") was labeled. Thus, an infant who is biased to look at a particular picture within a pair, regardless of the label presented, should end up with a Yoked score of zero for that pair, and Yoked scores above zero indicate lexical recognition.

Bergelson & Swingley (2012) found evidence of successful recognition of common words using both this yoked measures and the Accuracy measures we reported in the main results. We repeated the main analyses using the yoked scores (computed using performance in the Early window) as our measure of Visit 1 LPE to determine whether the results with these two measures of LPE converged. As reported below, we found no substantive differences in the results of analyses using the Yoked scores vs. Accuracy scores for the same time window.

Using one sample *t*-tests comparing the Accuracy scores to chance (i.e., to 0) We found that infants showed evidence of recognition for Easy words ($M = .07$, $SE = .028$; $t(64) = 2.41$, $p < .05$), but not for Hard words ($M = .03$, $SE = .032$; $t(62) = .90$, $p < .1$). Performance on the Easy trials was also correlated with receptive vocabulary percentile scores ($r(64) = .312$, $p < .05$), but performance on Hard trials was not ($r(65) = .06$, $p > .1$). Infants in the Large vocabulary group showed evidence of recognizing Easy words ($M = .12$, $SE = .035$; $t(32) = 3.67$, $p = .001$), but those with Small vocabularies did not ($M = .001$, $SE = .040$; $t(31) = .04$, $p > .1$). Neither infants in the Large nor Small vocabulary groups showed evidence of recognizing the Hard words (*t*s < 1.55, *p*s > .1). Better yoked scores on Easy trials also predicted faster RT at visit 2 ($r(32) = -.462$, $p < .01$), greater productive vocabulary size percentile scores ($r(43) = .388$, $p < .01$) as well as greater increases in productive vocabulary percentile scores across visits (r (43) = .355, p < .05). Using the Yoked scores, infants' performance on the Hard trials was not related to any of these measures (*r*s < .2, *p*s > .1).

In sum, we found highly convergent results when using the Yoked measure to assess Visit 1 LPE as when using Accuracy: Infants showed evidence of comprehending Easy words, and performance was related to concurrent normed receptive vocabulary size scores for both Accuracy and Yoked LPE measures.

The Easy Yoked scores also predicted visit 2 LPE RT, and changes in productive vocabulary size percentile scores.