Although the size of a child’s vocabulary associates with language-processing skills, little is understood regarding how this relation emerges. This investigation asks whether and how the structure of vocabulary knowledge affects language processing in English-learning 24-month-old children ($N = 32$; 18 F, 14 M). Parental vocabulary report was used to calculate semantic density in several early-acquired semantic categories. Performance on two language-processing tasks (lexical recognition and sentence processing) was compared as a function of semantic density. In both tasks, real-time comprehension was facilitated for higher density items, whereas lower density items experienced more interference. The findings indicate that language-processing skills develop heterogeneously and are influenced by the semantic network surrounding a known word.

There is wide variability in vocabulary growth across infancy, and these early differences can have important consequences for later cognitive and academic outcomes (Cunningham & Stanovich, 1997; Hart & Risley, 1995; Ramey & Ramey, 2004). Therefore, the factors that mediate the growth of early vocabulary skills are a subject of intense study. We explore novel connections between two factors—processing skill and conceptual development—which have been independently studied in their role in vocabulary growth in infancy. Specifically, we ask whether and how these two factors may be inter related by measuring the real-time recognition of words that vary in semantic density in 2-year-old children. Such a connection can extend theoretical and practical insight into the processes that underlie early vocabulary development and on the specific ways in which vocabulary development can support overall language processing. There are strong theoretical, empirical, and computational reasons to suspect that language-processing skills and conceptual development could be directly connected and dynamically interactive across early language development. We begin by reviewing how these conceptual development and language-processing skills are each independently related to vocabulary growth.

**Explaining Vocabulary Growth: Conceptual Development and Processing Skill**

Gopnik and Meltzoff’s (1987) “specificity hypothesis” was an early proposal that linked conceptual development and vocabulary growth. This idea posited a causal link between the onset of the vocabulary spurt and the understanding that objects can be semantically grouped into basic level categories. The proposal indicated that shifting a child’s conceptual organization could promote vocabulary growth by allowing the child to recognize semantic links between novel and known objects. This hypothesis was supported by numerous studies that linked success in exhaustive two-group sorting tasks with the onset of the vocabulary spurt in infancy (Gopnik & Choi, 1990; Gopnik, Choi, & Baumberger, 1996; Gopnik & Meltzoff, 1992; Mervis & Bertrand, 1995; Poulin-Dubois, Graham, & Sippola, 1995; but cf.
Are Language-Processing Skills and Semantic Development Inter-Related?

There has been little work that explores whether and how real-time language comprehension and conceptual development may be directly interconnected in infancy. Establishing this potential connection is the major aim of the current research. If conceptual development and processing skills are indeed linked, then we also expect the microstructure of vocabulary to associate with differences in linguistic processing within the same individual. This hypothesis leads to a testable prediction that the time course of recognition for a particular item should vary according to semantic network that surrounds the item. More specifically, we would expect facilitated processing for lexical items that have a relatively denser “semantic network” than those with fewer semantic neighbors. We test this prediction in the current study by asking whether and how semantic density directly influences processing skills during two language-processing tasks: lexical recognition and sentence processing (Altmann & Kamide, 1999; Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998).

The two selected language-processing tasks have been associated with vocabulary size in infants, such that higher overall vocabulary skill leads to facilitated performance, as measured by speed and accuracy in looking toward a target item (Fernald et al., 2006; Mani & Huettig, 2012). Yet, it is not known whether processing skill is tied directly to the semantic network that surrounds a lexical item or is simply driven by overall vocabulary size, as suggested by prior work. We describe assessment of semantic density and the experimental tasks in the following sections.

Defining Semantic Density

Although there are a number of reasonable potential metrics of semantic density that have been defined for the average adult lexicon (e.g., semantic connectivity, semantic set size, number of features; Mirman & Magnuson, 2008), the tremendous differences in early lexical knowledge between individual children present challenges for developmental research. However, because vocabulary in infancy is small, it is possible to take a comprehensive snapshot of the child’s productive vocabulary to measure semantic density within an individual child.

In the current study, we measure overall vocabulary with a common infant vocabulary instrument, the MacArthur-Bates Communicative Development Inventories (MBCDI; Fenson, 2007). The items in the Words and Sentences form of the MBCDI were selected to reflect many words that typically appear in infant productive vocabularies (Nelson, 1973). Moreover, nouns in this inventory are organized
according to several early-acquired semantic domains, such as animals, clothes, body parts, and vehicles. These aspects of the MBCDI, therefore, have the potential to yield meaningful measures of semantic density within early-acquired categories. For the current study, we operationalize semantic density for each lexical item as the proportion of words that child says within its corresponding semantic category. Concretely, the semantic density of *dog* would be defined as the proportion of all possible *animal* words on the MBCDI that the child says. We use items from six early-acquired semantic domains including animals, body parts, clothes, drinks, fruits, and vehicles.

Using Eye Tracking to Measure Processing Skill in Infants

Measurements of gaze in response to spoken language are often used as an index of language-processing skill in infancy (e.g., Fernald, Zangl, Portillo, & Marchman, 2008). This technique, variously termed the “looking while listening” or “visual world paradigm,” measures changes in gaze in relation to the unfolding of speech in real time (Huettig, Rommers, & Meyer, 2011). For example, Fernald et al. (1998) presented children with arrays containing a distractor and target image (e.g., car and shoe) while simultaneously labeling one image (*Look at the shoe!*), and found that the speed and accuracy to view the correct target item improved from 15 to 24 months.

Semantic Density and Processing Speed: Facilitation or Interference?

Prior findings (reviewed earlier) reveal that larger vocabulary size boosts lexical and sentence recognition (Fernald et al., 2006; Mani & Huettig, 2012). By extension, these findings support the possibility that semantic density should facilitate lexical recognition. Contrasting evidence from the phonological density literature indicates that dense networks may instead promote lexical interference. In this case, increasing semantic density may result in enhanced competition from semantically similar items in the child’s lexicon, resulting in slower lexical recognition in denser semantic networks. Such a pattern would be analogous to inhibitory effects on word recognition in young children and adults in words occurring in dense phonological neighborhoods (Garlock, Walley, & Metsala, 2001; Mani & Plunkett, 2011; Vitevitch, Luce, Pisoni, & Auer, 1999).

Although there is no direct evidence that young children experience enhanced semantic interference in dense networks, recognition of semantic and phonological links between words does change considerably over the first 2 years of life. Notably, semantic priming effects have been inconsistently observed at 18 months and reliably reported by 24 months in a variety of methods (Arias-Trejo & Plunkett, 2010; Rämä, Sirri, & Serres, 2013; Styles and Plunkett, 2009; Torkildsen, Syversen, Simonsen, Moen, & Lindgren, 2007; Willits, Wojcik, Seidenberg, & Saffran, 2013). In contrast, although phonological priming appears to be present at 18 months (Mani & Plunkett, 2010), this pattern shifts from facilitation to interference by age 2 (Mani & Plunkett, 2011). This shift may be due to the changing nature of vocabulary density and structure. Mayor and Plunkett (2014a) carried out simulations that suggest that inhibitory (phonological) connections between words increases as vocabulary size (and density) increases. It is not clear if the semantic structure of the early lexicon similarly influences mechanisms of lexical activation and recognition.

Therefore, we test 24-month-old children on two language-processing tasks that are sensitive to vocabulary effects: (a) a lexical-recognition task and (b) a simple sentence-processing task consisting of a semantically informative verb followed by a related thematic object. As described earlier, children’s individual knowledge with six early-acquired domains is also assessed through parental report on the words and sentences form of the MBCDI. In both studies, we posit that if semantic density facilitates lexical and sentence processing, then we would expect to see relatively more robust recognition of higher (vs. lower) density target items. Alternatively, if semantic density instead creates greater lexical competition, then we should expect to find the opposite pattern, with boosted recognition for low- (vs. high-) density items.

Method

Participants

Thirty-two 24-month-old English-learning children (18 F, 14 M) from the metropolitan region surrounding a large city of southern California were recruited via advertisements in the community. All participants had previously participated in a separate experimental task and cognitive screening using the Bayley Scales of Infant Development, 3rd ed., Cognitive Subscale (BSID-III; Bayley, 2006) at 18 months of age. All infants received a standard BSID-III score of 85 or above (*M* = 106.3, *SD* = 11.9, range = 85–130), meaning that no child fell more
than 1 SD below the mean in cognitive ability. Data collection took place between November 2012 and September 2013.

The demographic characteristics of our sample were relatively diverse, with 39% reporting membership in an ethnic minority group. Our families were highly educated. All mothers had completed at least high school, and 82% had graduated from college. Additionally, 89% of children were living in two-parent households. Infants were learning English (hearing fewer than 20 hours per week of another language at home), and parents reported that their children had normal vision and hearing and no other concerns about their child’s language or cognitive development. All participants had normal birth histories and no recent or chronic ear infections.

Stimuli

Selection of Category Domains and Items

The MBCDI checklist is organized into several category sections that correspond to early-acquired vocabulary items. We consulted the Cross-Linguistic Lexical Norms database (CLEX; Dale & Fenson, 1996) to select six category domains with wide representation in 24-month-old vocabularies, including animals, clothing, vehicles, body parts, and two subcategories from food, fruits and drinks. More information about the categories can be found in Table 1.

We used the CLEX database (http://www.cdiclex.org) to select two highly known words in each category (i.e., produced by at least 70% of American 2-year-olds), with one word in each category assigned to the lexical-recognition task, and the other to the sentence-processing task. The items in the lexical-recognition task were: animals: bird; fruit: banana; clothing: diaper; vehicles: airplane; drinks: juice; body parts: nose. The items for the sentence-processing task were: animals: dog; fruit: apple; clothing: shoe; vehicles: car; drink: milk; body parts: teeth.

Visual Stimuli

We selected a colorful, photorealistic 400 x 400 pixel image for each experimental item centered on a white background. Several additional images were used either to direct the child’s attention toward the screen during the study (e.g., a small flower, or happy face) or to maintain infant interest (e.g., a large image of the Sesame Street character, Elmo).

Auditory Stimuli

Speech stimuli for all items were recorded in an infant-directed voice by a female native speaker of California American English (E.E.) on a mono channel at 44,100 kHz sampling rate. To precisely control the onset and offset of the auditory stimuli, the duration of all linguistic stimuli in both tasks was calculated in Praat (Boersma & Weenink, 2012).

In the lexical-recognition task, the auditory items were the spoken words that were selected for the study. The lengths of these words were automatically adjusted in Praat to a mean length of 1,020 ms and intensity of 70 dB. In the sentence-processing task, the linguistic stimulus consisted of a simple sentence beginning with a semantically selective verb, followed by a noun from one of the selected semantic domains. The sentences in each

<table>
<thead>
<tr>
<th>Category</th>
<th>Total MBCDI items</th>
<th>High % of children</th>
<th>High N</th>
<th>Low % of children</th>
<th>Low N</th>
<th>Mean proportion (SD)</th>
<th>Range (median)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animals</td>
<td>43</td>
<td>.687</td>
<td>22</td>
<td>.313</td>
<td>10</td>
<td>.745 (.225)</td>
<td>17–43 (33.5)</td>
</tr>
<tr>
<td>Body parts</td>
<td>27</td>
<td>.687</td>
<td>22</td>
<td>.313</td>
<td>10</td>
<td>.795 (.240)</td>
<td>2–27 (24.9)</td>
</tr>
<tr>
<td>Clothing</td>
<td>28</td>
<td>.094</td>
<td>3</td>
<td>.906</td>
<td>29</td>
<td>.623 (.221)</td>
<td>5–26 (18.0)</td>
</tr>
<tr>
<td>Drinks</td>
<td>7</td>
<td>.281</td>
<td>9</td>
<td>.719</td>
<td>23</td>
<td>.664 (.175)</td>
<td>3–7 (6.2)</td>
</tr>
<tr>
<td>Fruits</td>
<td>7</td>
<td>.656</td>
<td>21</td>
<td>.343</td>
<td>11</td>
<td>.769 (.220)</td>
<td>3–7 (6.0)</td>
</tr>
</tbody>
</table>

Note. MBCDI = MacArthur-Bates Communicative Development Inventories.
domain were: animals: pet the doggy; body parts: brush the teeth; clothing: wear the shoe; drink: drink the milk; fruit: eat the apple; vehicles: drive the car. As with the lexical-recognition task, all sentences were normalized for duration and intensity such that the duration of each word in the sentences was identical across items. The total sentence durations were 1,449 ms, with verb duration of 563 ms, article duration of 189 ms, and noun duration of 697 ms.

The speaker also recorded several encouraging items and tag phrases that were intended to maintain the infants’ attention during the study. For example, each lexical item was followed by a tag phrase such as “That’s cool!” or “Do you like it?” and the lexical item was preceded by the speaker saying, “Look!” Other encouraging phrases were presented in between experimental trials along with other colorful images. All items were normalized to 70 dB intensity.

Procedure

Vocabulary Assessment

Parents were mailed a MBCDI words and sentences form approximately 1 week before their laboratory visit and returned the completed checklist during their appointment. In addition to the MBCDI, parents completed an additional comprehension checklist for all items that appeared in the study on the day of their laboratory visit. This questionnaire asked parents to mark whether their child understood the words to be used in the study, including all nouns and verbs in both tasks. Parents rated comprehension for each item on a scale of 1 = child does not understand the word to 4 = child definitely understands the word. Because we were only interested in toddlers’ responses to known words, we removed experimental responses to any items where their parents marked a “1” or “2” to the target item.

Measurement of Lexical Recognition and Sentence Processing

Infants were seated either in a car seat or on their caregiver’s lap while they viewed images on a 17-in. LCD monitor positioned by a flexible arm mount. The eye tracker camera was mounted directly underneath the monitor and was positioned between 580 and 620 mm from the infant’s forehead. Auditory stimuli were presented via a speaker placed behind the monitor. During the experiment, caregivers listened to music via head-phones and were instructed not to name the images.

Before the experimental trials began, the position of the eye tracker was adjusted and the camera was focused while the infant viewed a short video. The tracker was then calibrated with a 5-point routine using an animated looming bull’s eye image paired with a whistling sound. Toddlers generally viewed these calibration images without any explicit instruction, but the experimenter pointed to the screen when necessary to help direct the child’s gaze toward the appropriate images.

After the tracker was calibrated, the experimental trials began. Each trial began with the presentation of central colorful 30 x 30 pixel image. Once the toddler fixated toward this image, it disappeared and the target and distractor image were presented on the left and right side of the screen for 2,000 ms in silence. The silent preview period served to familiarize the child with the object images and their locations in advance of the target label. Next, a 100 x 100-pixel image (e.g., a smiling sun, a star) appeared at the center of the screen as the sound “Look” was played. The central image automatically remained on the screen until the child fixated within its bounds for 100 ms. Note, our stimulus arrangement placed the edges of central stimulus and target/distractor images approximately 1.3° visual angle apart, and the center of the stimulus within 5° of visual angle. This visual angle value is within the 5°–10° of useful visual field that is commonly noted within visual search studies that include multiple objects (reviewed in Irwin, 2013). Our stimulus arrangement therefore ensured that each object remained within a useful field of view (i.e., requiring minimal working memory commitment for object location), despite drawing momentary attention to the center stimulus. We implemented this gaze-contingent procedure to ensure that the infant was attentive to the screen before proceeding with the spoken stimulus. Once the infant fixated to the center image, the auditory stimulus was spoken (e.g., “Car!” or “Pet the doggy!”) followed by an encouraging tag phrase (e.g., “That’s cool!”) and the central image disappeared. The target and distractor images remained on the screen for 4,000 ms post label onset (see Figure 1, for an illustration of the task).

Each image appeared four times across the course of the study, as a yoked pair with another object. The yoked object pairs in the lexical-recognition task were banana/ juice, diaper/nose, and
bird/airplane. In the sentence-processing task, yoked pairs were presented across the same categories with the following items: apple/milk, dog/car, and shoe/teeth. Each object was twice labeled as the target and twice unlabeled as the distractor image (and the order of target and distractor presentation varied across versions), and each image appeared on the left and right pseudorandomly with equal frequency across these four trials. This yielded 12 trials in each task (24 total trials) that were interspersed and distributed equally across three blocks. Filler trials appeared between each block with images containing complex scenes and exciting characters (e.g., Elmo), along with encouraging statements such as “Wow, you’re doing great!” Within any one version, the target and distractor images were equally likely to appear on the left and right side across all categories. The entire procedure lasted approximately 5–10 min.

Recording of Eye Movement Data

Moment-by-moment visual behavior was recorded from image onset to offset at 500 Hz by an SR Research EyeLink 1000 (SR Research: Ottawa, Ontario, Canada) eye tracker and binned into 50 ms intervals for offline analysis. Eye movements were automatically classified into fixations, saccades, and blinks according to default tracker settings and areas of interest were defined as the two 400 × 400 pixel locations of the target and distractor images.

Approach to Analysis

Assignment of High- and Low-Category Domains

Because our primary interest was to measure how semantic density influences the recognition of that item, independent of the effects of vocabulary size, we implemented a metric of relative (rather than absolute) semantic density with respect to each child’s vocabulary. For each category, we calculated each child’s semantic density by dividing the number of words that the child says in the category by the total number of words surveyed in each category domain. Then, for each child, the categories with the three highest proportions were assigned to the high-density condition, and the lowest proportion categories were assigned to the low-density set. We use a median-split assignment for high density and low density rather than a continuous proportion measure to control for overall vocabulary size across participants. In three participants, the third- and fourth-ranked category proportion was equivalent, so we assigned both categories to the high-density condition. This procedure therefore yielded a unique arrangement of experimental items that was classified as high density and low density for each infant. The distribution of category
In both tasks, we explored how looks toward the labeled target item vary as a function of domain knowledge over larger and smaller windows of analysis. For the larger windows of analysis, we initially calculated a measure of accuracy that is commonly used in infant eye-tracking research, which is defined as the number of fixations toward the target divided by the sum of fixations toward the target and distractor images. With this metric, accuracy ranges from 0 to 1, with values > 0.5 indicating a preference to look toward the target (vs. distractor) over the specified time region. In the lexical-recognition task, this accuracy metric was averaged over a broad time window from 300 to 1,800 ms post label onset to correspond with time windows used in other studies of infant lexical recognition (Fernald et al., 2008). In the sentence-processing trials, there were two potentially informative events of interest across the linguistic stimulus (the spoken verb and the sentence final object). Thus, we defined two time windows of interest for statistical analysis, following precedent from prior research using a similarly structured task (Mani & Huettig, 2012). The first time window, the anticipatory window, was defined as the portion of the sentence that occurred before the onset of the sentence final noun, spanning from 300 post verb onset to the 750 ms time bin (noun onset was 752 ms). The second time window, the noun window, spanned from 300 ms post noun onset to the 1,800 ms post noun onset (i.e., from the time bins spanning 1,050 to 2,550 ms post sentence onset).

Next, we carried out an finer grained analysis of real-time comprehension in both tasks using smaller 50 ms time bins to address whether and when there were differences in looks to the target (vs. distractor) as a function of semantic density. We adopted a dependent measure, log-gaze proportion ratio, that is centered around zero and varies between positive and negative infinity. With this measure, larger scores indicate a relative advantage to view the target (vs. distractor). Log-gaze ratios have been adopted in recent years in the eye-tracking literature as a way to overcome violations of statistical assumptions of linear independence (because increases in fixations to one item necessarily decrease looks to the alternate image) and homogeneity of variance (because simple proportion ratios vary between 0 and 1; see Arai, van Gompel, & Scheepers, 2007; Borovsky, Sweeney, Elman, & Fernald, 2014; Knoeferle & Kreysa, 2012, for a similar approach). In each time bin, we calculated the log-gaze proportion ratio for the target versus distractor as \( \log(p_{\text{target}}/p_{\text{distractor}}) \). Because log ratios are undefined for zero values, we replaced every instance of a zero value in the numerator or denominator with a value of 0.01.

Using this log-gaze metric across 50 ms time bins, we addressed our primary questions regarding differences in domain density in lexical-recognition and sentence-processing trials using a statistical approach that seeks to detect reliable differences across time between conditions while controlling for multiple comparisons (Type I error). We implemented a nonparametric cluster-based permutation analysis that has been most commonly applied to FMRI (functional magnetic resonance imaging) and ERP (event-related potentials) time series analyses (Groppe, Urbach, & Kutas, 2011a; Maris & Oostenveld, 2007) and has become more common in eye-tracking data analysis (Barr, Jackson, & Phillips, 2014; Von Holzen & Mani, 2012). This test calculates a cluster \( t \) statistic that sums across temporally adjacent point-wise \( t \) values that exceed a predefined threshold. This cluster \( t \) statistic is then compared to a null-hypothesis distribution of cluster \( t \) values that are generated via a Monte Carlo permutation approach (outlined in Barr et al., 2014, appendix). Following standards set by Groppe, Urbach, and Kutas (2011b), we used 2000 random permutations to generate a distribution of the null hypothesis with sufficient precision to control family-wise error rate to \( < 0.05 \).

We used this cluster approach in both tasks to address whether differences in the time course of linguistic processing vary by domain knowledge by directly comparing high- versus low-density log-gaze proportion ratios across time (using a threshold of \( p < .05 \), two sided for point-wise comparisons). We additionally asked when fixations to the target exceed those of the distractor by identifying time clusters where log-gaze proportions exceed 0 (using a threshold of \( p < .05 \), one sided) for high- and low-density domain items, separately.

Because we are primarily interested in how responses to known words are influenced by the structure of the child’s vocabulary, we only included trials in analysis where the parent reported that their child understood the target word in both tasks, as well as the verb in the sentence-recognition task. Using this criterion, we excluded two trials (of 384, or 0.52% of total trials) in the lexical-recognition task and another eight
trials (of 384, or 2.1% of total trials) in the sentence-processing task. Next, we removed trials where infants focused on the target and distractor for < 20% of the 300–1,800 ms analysis time window for lexical-recognition trials, or < 20% of the 300–2,550 analysis windows for sentence-processing trials. We adopted this threshold following precedent set in other eye-tracking studies with young children, such as Nordmeyer and Frank (2014), who removed trials with over 30% of samples missing, and Quam and Swingley (2014), who excluded trials where children fail to view the pictures for 17% of the time window (for 300 ms out of a 1,650 ms analysis window).

This 20% criterion led us to remove 36 of the remaining 382 trials in the lexical-recognition task, or 9.4%, and an additional 27 trials of the remaining 376 trials in the sentence-processing task (7.2% of remaining trials). In total, 38 trials were excluded from analysis, or 9.9%, of the lexical-recognition task data set, leaving 346 remaining trials in the final analysis. In the sentence-processing data set, 8.9% of total trials were removed, leaving 350 trials over which subsequent analyses were performed.

Results

Lexical-Recognition Task

Figure 2A illustrates the time course of fixation proportions toward the target and distractor images for high- and low-density items in the lexical-recognition task. As expected, there was a rapid rise in fixations post label onset and fixations toward the target appeared to exceed those of the distractor within 800–1,000 ms post label onset in both high- and low-density conditions before the offset of the label at 1,020 ms. The time course plot also illustrates differences between the real-time recognition of high- and low-density items. Infants appear to be faster to fixate uniquely toward the target item in the low- versus high-density condition. This difference appears to be partially driven by a rise in fixations toward the distractor item relative the target in the low-density condition that does not exist in the high-density condition. This pattern is echoed in the log-gaze ratio plots in Figure 2B, which indicate an early advantage in distractor versus target fixations (negative scores) for low- but not high-density items.

We initially carried out an analysis of target accuracy from 300 to 1,800 ms across high- and low-density items (Figure 3). Target accuracy exceeded 0.5 in both conditions, $t_{high}(31) = 7.5$, $p_{high} < .0001$, $d_{high} = 2.69$; $t_{low}(31) = 7.09$, $p_{low} < .0001$, $d_{low} = 2.54$, indicating that participants successfully interpreted the label in this time window in both conditions. Further analysis also found no significant differences as a function of semantic density ($M_{high} = 0.68$, $SD_{high} = 0.13$; $M_{low} = 0.65$, $SD_{low} = 0.12$), $t(31) = 0.96$, $p = .34$, $d = 0.34$. These comparisons indicate that listeners identified the appropriate
target item in both high- and low-density conditions, but there were no differences due to density in overall accuracy of gaze toward the target item over this relatively broad time window (Figure 3).

The next analyses addressed whether and when there were finer grained differences across the time course of lexical recognition as a function of domain density. We first asked when looks to the target exceeded those to the distractor (indicating recognition of the label) with a cluster-based permutation analysis that identified the time bins when log-gaze proportion ratios were significantly positive for high- and low-density items. This analysis indicated a target preference from 650 to 1,800 ms for high-density items (cluster $t = 130.8$, Monte Carlo $p = .002$) and from 850 to 1,800 ms for low-density items (cluster $t = 124.2$, Monte Carlo $p = .001$). Together, these results indicate that participants rapidly interpreted the spoken label, by directing their fixations consistently toward the target item before the label was completely spoken (label offset was at 1,020 ms) and continuing to do so until the end of the analysis window. Additionally, this analysis suggests that the identification of the target item was approximately 200 ms faster in the high- versus low-density domain.

We next directly compared log-gaze proportions in 50 ms time bins across high- and low-density domains using a cluster-based permutation analysis, described earlier. This direct comparison of domain effects indicated a difference between domains from 300 to 650 ms post label onset (cluster $t$ statistic = 18.32, Monte Carlo $p = .027$). This difference, as seen in Figure 2B, was driven by relatively greater log-gaze ratios for the high- (vs. low-) density items and supports the findings in the first cluster analysis that participants were relatively faster to settle on the appropriate target item in the high-density condition.

The Role of Distractor Semantic Density

As one reviewer helpfully noted, the semantic density of the distractor object may exert an independent influence on target recognition. We therefore carried out an exploratory analysis of distractor density on target recognition in four main conditions: (a) high-density target–high-density distractor (high T–high D), (b) high-density target–low-density distractor (high T–low D), (c) low-density target–high-density distractor (low T–high D), and (d) low-density target–low-density distractor (low T–low D). We note that our experimental trials were not balanced with respect to distractor. Therefore, participants did not contribute trials equally across all conditions in this analysis. Total trials in each condition were high T–low T: 75, high T–low D: 100, low T–high T: 99, and low T–low D: 72. We conducted a $(2 \times 2)$ repeated measures analysis of variance (ANOVA) with target density (high and low density) and distractor density (high and low density) as factors. There were no main effects of target density, $F(1, 70.29) = 0.89, p = .35$, and distractor density, $F(1, 70.29) = 0.07, p = .78$, and there was no significant interaction effect, $F(1, 88.9) = 1.73, p = .19$. Follow-up Tukey analyses did not reveal significant differences among any of the individual comparisons.

Sentence-Processing Task

The time course of fixations toward target and distractor images for high- and low-density items in the sentence-processing task is illustrated in Figure 4A. As expected, we saw a rapid rise in fixations toward the target object before it was mentioned, consistent with prior literature on prediction that suggests that semantically selective verbs facilitate (anticipatory) prediction of thematic objects (Altmann & Kamide, 1999; Mani & Huettig, 2012). We initially carried out an analysis of target accuracy across two broad time windows: (a) across the anticipatory period (300–750 ms post sentence onset), and (b) during the noun time window (1,050–2,550 ms post sentence onset), illustrated in
In the anticipatory window, accuracy for the high-density items ($M_{\text{high}} = 0.54, SD_{\text{high}} = 0.20$) exceeded that of the low-density items ($M_{\text{low}} = 0.42, SD_{\text{low}} = 0.24$), $t(31) = 1.99, p = .056, d = 0.71$, though neither value significantly exceeded 0.5, $t_{\text{high}}(31) = 1.13, p_{\text{high}} = .13, d_{\text{high}} = 0.41; t_{\text{low}}(31) = -1.98, p_{\text{low}} = .97, d_{\text{low}} = -0.71$. This comparison indicates that there was a difference as a function of domain density in recognizing the coordinating verb, although there is not strong evidence for prediction of the unspoken noun in this time window.

In the noun window, accuracy did not differ as a function of domain density ($M_{\text{high}} = 0.76, SD_{\text{high}} = 0.13; M_{\text{low}} = 0.73, SD_{\text{low}} = 0.13$), $t(31) = 0.75, p = .46, d = 0.27$, and accuracy exceeded 0.5 for both high-density, $t(31) = 11.01, p < .001, d = 3.95$, and low-density items, $t(31) = 10.38, p < .001, d = 3.73$. These comparisons indicate that listeners viewed the target item in response to the noun label in high- and low-density conditions, but there were no differences due to density in overall accuracy of gaze toward the target item over this broad time window (Figure 5).

We next carried out a fine-grained comparison of log-gaze fixation ratios for high- and low-density items across the time course to determine whether there is a difference in the timing of fixations according to domain knowledge. In the anticipatory window, a cluster-based permutation analysis did not identify any periods of time that significantly exceeded 0 in the high- and low-density conditions. However, there is a significant density difference from 300 to 650 ms post verb onset (cluster $t$ statistic = 17.04, Monte Carlo $p = .032$). This difference was driven by relatively larger log-gaze ratios (i.e., a preference to view the target) for high- versus low-density items in the anticipatory window. In the noun window, log-gaze ratios significantly exceeded 0 across the entire time window for both high-density (cluster $t = 215.5$, Monte Carlo $p = .002$) and low-density (cluster $t = 187.3$, Monte Carlo $p = .002$) items. There were no significant
cluster differences across density in the noun window (cluster t statistic = 7.03, Monte Carlo p = .18).

Together, the findings across both windows indicate that participants recognized the target once it was labeled, in both high- and low-density conditions. Importantly, differences due to density emerged during the anticipatory window before the target was labeled. Because log-gaze ratios did not significantly exceed 0 during this period, this analysis indicates that density effects in the anticipatory window were not driven by differences in viewing the distractor image. Children were more likely to view the distractor for low-density items. This pattern suggests that density effects in the anticipatory window were driven by greater interference from the distractor item in low-density domains.

The Role of Distractor Semantic Density

As in the lexical-recognition task, we explored the effect of distractor density on the recognition of the target word across high- and low-density target and distractor conditions with a 2 × 2 ANOVA with target density (high and low) and distractor density (high and low) as factors in both the anticipatory time window and the noun window. As stated earlier, trials were not balanced across conditions, such that the total number of trials in each condition was high T–high D: 77, high T–low D: 99, low T–high D: 102, and low T–low D: 72. In the anticipatory window there was a main effect of target density, \( F(1, 95) = 7.89, p = .006 \), with greater accuracy for high-density targets. There was also an effect of distractor density, \( F(1, 95) = 6.63, p = .012 \), with greater looking toward the target in the presence of a high-(vs. low-) density distractor. There were no significant main or interaction effects in the noun window (all ps > .2). This analysis indicates that, during the anticipatory window, 2-year-olds showed greater looking toward the target when it was from a high-density domain, as well as when the distractor item was also from a high-density domain. Although this pattern might reflect greater referential certainty and anticipatory processing in conditions where both the target and distractor images are from categories that have relatively greater semantic density, this effect did not carry over into time period when the target object was labeled.

Discussion

Whereas previous studies had found that vocabulary growth was associated with linguistic-processing skill and semantic development individually, little was known regarding the relation between processing and semantic development. We therefore asked whether real-time lexical recognition in infancy varied according to the structure of the 24-month-olds’ semantic knowledge using a within-subjects experimental design.

We outlined two possible outcomes. One hypothesis (the interference hypothesis) predicted that increased semantic density would boost semantic competition from related items and result in less robust lexical recognition for high- (vs. low-) density items. This pattern would be analogous to findings where lexical recognition is slowed for items with higher phonological neighborhood density (Garlock et al., 2001; Vitevitch et al., 1999). Another plausible but contrasting potential outcome (the facilitation hypothesis) was that greater semantic density within a category might boost lexical recognition, as having knowledge of many semantically related items could increase activation for the intended lexical items. This outcome is predicted, for example, by distributional and semantic feature models of lexical activation, where activation for any item can be facilitated by activation of similar meanings that share many features or contexts (Cree, McRae, & McNorgan, 1999; Landauer & Dumais, 1997; Plaut & Booth, 2000).

Our findings are most consistent with the latter potential outcome. In both experiments, fixations to higher density targets exceeded lower density target fixations, as predicted by the lexical facilitation hypothesis. However, this finding was partially driven by an early preference for the distractor item in low-knowledge conditions. This early interference pattern suggests that low-knowledge items experienced semantic interference from distractor items, whereas high-knowledge items did not. This pattern suggests that recognition of the high-density items may not be purely driven by boosted lexical activation for high-knowledge items insomuch as it is simultaneously driven by reduced semantic interference.

We should also note that this distractor preference existed despite the fact that our counterbalancing scheme ensured that all items were equally likely to serve as targets and distractors, thereby ruling out a simple explanation of this effect due to visual saliency or preference for some items over others. Similarly, although we used the same categories across tasks, we selected different (highly frequent) items that are commonly known to 18- to 24-month-old children in each task (Dale & Fenson, 1996). Thus, this distractor preference for low-
density conditions remained consistent across different items. We also individually verified knowledge of the experimental items (nouns and verbs) via parental report. Thus, the density effects are unlikely to be driven by a failure to understand the experimental items. In sum, it appears lexical and sentential processing varies as a function of semantic knowledge, such that items that have a denser network of semantically related items experience greater simultaneous lexical facilitation and reduced interference than items that have sparser semantic category neighborhoods.

Perhaps one of our more intriguing findings is that semantic density influenced performance in a sentence-processing task. In this case, we found no strong evidence for linguistic prediction per se nor was there an effect of density as the target image was labeled. Instead, density effects emerged during the anticipatory window before the target noun was even mentioned. Additionally, the distractor image also contributed to this effect. We discuss each of these points individually below.

With respect to the lack of strong support for anticipatory processing, several points about our design are relevant here. First, we do not include a semantically neutral verb condition unlike other paradigms that explore prediction with semantically selective verbs. The main reason for this decision is that our design is not necessarily about prediction per se but rather an extension of how semantic density may influence the linguistic processing beyond a single named object. Second, the timing of our sentence materials was faster than that of similar previous studies. For example, the stimuli in Mani and Huettig (2012) had at least 1,500 ms from the onset of the sentential verb and disambiguating noun, whereas the same window in the current study was 752 ms in duration. It seems plausible that these factors may explain the lack of strong anticipatory effects in our task. Future work is needed to address how speech rate contributes to predictive processing in young children.

Another perplexing finding from our sentence task was that there was not an effect of density as the sentence final object was spoken. This pattern differs from the lexical-recognition task, where recognition of a spoken noun (which was not preceded by a semantically constraining verb) varied as a function of semantic density. Instead, density effects appeared before noun onset, as the verb was spoken during the anticipatory window. It seems likely that semantic density effects resolved before the onset of the noun, suggesting that 2-year-olds had already begun to (pre)activate coordinating semantic information associated with the target noun while the verb was spoken.

Next, with respect to our findings of density effects during the anticipatory window, we found that density differences arose as the semantically selective verb was spoken—preceding the onset of the target label. This difference was driven by a greater tendency for participants to view the distractor item during the anticipatory window for lower versus higher density items. This finding suggests that lower density items experienced greater lexical interference from competing distractors than did high-density items but only during the anticipatory window. This finding mirrors our results from Experiment 1, where lower density target items experienced relatively greater distractor interference at early points in linguistic processing as well.

At the same time, we also note an effect of the distractor image density on target recognition during the anticipatory window. In this case, target recognition was facilitated when the distractor image also belonged to a higher density domain. We interpret these findings with caution, because the experiment was not initially designed to carefully control for distractor effects across all participants. Our results nevertheless are suggestive that target recognition is facilitated by greater semantic density, and this effect extends from the semantic neighborhood surrounding both the target and distractor items. In this case, it appears that target recognition is boosted in conditions where there is greater referential certainty about the distractor object so that it may more easily be discounted as a potential lexical referent.

Taken together, the sentence-processing task findings indicate that semantic domain knowledge at the noun level has a general effect on language processing that extends beyond the individual lexical item. In other words, verb recognition varies according to the density of near semantic neighbors using similar lexical activation mechanisms that exist for nouns.

We next comment more generally on the lexical interference effect that emerged for low-density items across both experimental tasks. This interference effect for low-density items was relatively unexpected given our knowledge of lexical interference from prior literature in young children. What are the mechanisms that might account for this pattern? One possibility is that items from lower density categories may, in general, show relatively less robust lexical activation than items in higher density domains. This difference may stem from somewhat weaker or less tightly connected semantic
representations for lower density items, which, in turn, would create increased interference from unrelated distractor items present in the array. Such an explanation seems consistent with prior studies that fail to find consistent semantic priming effects in 18-month-old children (e.g., Arias-Trejo & Plunkett, 2010; Styles and Plunkett, 2009). Younger children’s relatively smaller lexicons are likely to be less dense compared to older 24-month-old children. Clearly, it will be important to understand how the dynamics of lexico-semantic processing evolve in tandem with the growing lexicon.

In sum, our findings also add to growing evidence that early infant lexicons are organized semantically, much like adult lexicons. A number of prior studies now indicate that infants as young as 18–24 months recognize lexical similarities between known and novel words according to taxonomic, perceptual, thematic, and structural relations (Mani & Borovsky, in press). There is also some evidence that the recognition of these relations may change according to the overall vocabulary knowledge of the individual (Râmâ et al., 2013). Our findings additionally demonstrate that speech processing skills are tied not only to the overall knowledge of the child but vary according to individual knowledge and experience with items in different domains.

Most importantly, these results suggest that psycholinguistic mechanisms of lexical activation and recognition that might be assumed to operate homogeneously across our entire lexicon instead develop idiosyncratically and heterogeneously according to the child’s knowledge and experience. More generally, our research supports a growing body of evidence that early experiences can have a tremendous impact on categorical and lexical learning. For example, early home experiences (e.g., experience with pets) can influence learning and attention to a category (e.g., cats) as early as 4 months of age (Hurley & Oakes, 2015; Kovack-Lesh, McMurray, & Oakes, 2014). Mayor and Plunkett (2014b) have also recently described a U-shaped pattern in early variability in word knowledge between children, with expressive vocabulary increasing in variability between 15 and 24 months of age, before becoming more coherent between children. Together, these findings indicate that early experiences may have a particularly substantial impact on early learning and attentional mechanisms, primarily in the first 2 years, and that there is a need to explore whether and how early differences in knowledge and experience may lead to lasting impacts in the developmental trajectory and outcomes on early lexical and cognitive skills.

These findings also connect with a growing literature that explores how the current organization of the lexicon influences vocabulary growth across infancy. Rather than learning a random cohort of words, there is increasing evidence that children learn words in a “clustered” fashion, via a tendency to add new words to their vocabulary that are related to already known items (Beckage et al., 2011; Hills et al., 2009; Steyvers & Tenenbaum, 2005). This pattern of vocabulary growth, termed “preferential attachment” by network researchers (Barabási & Albert, 1999), and initially connected to semantic vocabulary growth by Steyvers and Tenenbaum (2005), may facilitate word learning by providing a basis by which learners generalize prior knowledge to novel words that share some semantic overlap (Borovsky & Elman, 2006). The structure of knowledge may also affect the basic strategies that children use to acquire words (Colunga & Sims, 2011; Yurovsky et al., 2012). Although the measures that we use are not identical to the previously mentioned network modeling analyses, our metrics of semantic density within categories do capture elements of semantic microstructure of early vocabulary. Importantly, this work suggests that increasing semantic density may simultaneously reduce potential interference from lexical competitors and boost activation for intended lexical items. These processes extend to novel word learning, where it is clear that 2-year-old children encode and recognize semantic relationships between novel items that share similar perceptual features and that appear in similar sentence contexts (Wojcik & Saffran, 2013; Wojcik & Saffran, 2015). Similarly, lexical processing is facilitated for novel items within denser semantic domains (Borovsky, Ellis, Evans, & Elman, in press). Additional work is needed to delineate how mechanisms of lexical competition and facilitation interact during known word recognition and novel word learning.

We also note some limitations in the current study. First, this study explores the effect of semantic density on processing at a single age. It will be important to extend this work to address whether and how domain knowledge influences performance as vocabulary size and structure changes across development. Another issue arises from the fact that some categories were, on average, more likely to fall in a high- or low-density domain than others (e.g., clothes). Future research would need to explore whether these overarching differences may drive density effects due to the salience or frequency in the input for these particular categories. A further limitation is that our measure of semantic
density relies on a measure of expressive vocabulary and not receptive knowledge. Recent work suggests that productive vocabulary may have a tighter relationship with processing skills than those of receptive vocabulary in 24-month-olds (Mani & Huetting, 2012); however, additional work is needed to explore whether this link also extends to semantic density as well.

Conclusions

Vocabulary growth is connected to the development of a plethora of linguistic and nonlinguistic skills. Ultimately, a full understanding of language acquisition will rest not only in identifying the relevant mediators of vocabulary growth but also in characterizing the interactions between these factors. Our studies provide initial evidence that the semantic organization of items within the child’s lexicon leads to important changes in how word meanings are activated and recognized. Our findings additionally reveal some initial clues as to how these processes may scale to complex multiword language-processing tasks. Our results additionally suggest that “language-processing skill” is not a purely endogenous or unitary construct as it varies within an individual according to their experience and knowledge. These findings indicate a hopeful possibility that organizing vocabulary training to items within semantic domains may yield benefits for processing skills and vocabulary growth. We plan to explore whether targeted vocabulary instruction within semantic domains could improve processing skills and, by extension, language and academic outcomes more generally.

References


