Published in final edited form as:
Dev Psychol. 2022 April ; 58(4): 607–630. doi:10.1037/dev0001291.

Lexico-Semantic Structure In Vocabulary And Its Links To Lexical Processing In Toddlerhood And Language Outcomes At Age Three

Arielle Borovsky
Department of Speech, Language, and Hearing Sciences Purdue University

Abstract

Toddlerhood is marked by advances in several lexico-semantic skills, including improvements in the size and structure of the lexicon and increased efficiency in lexical processing. This project seeks to delineate how early changes in vocabulary size and vocabulary structure support lexical processing (Experiment 1), and how these three skills together (vocabulary size, structure and lexical processing) relate to later language outcomes at age 3 (Experiment 2). Experiment 1 explored how the size and semantic structure of toddlers' vocabulary from 18 to 24 months (N=61) predicted performance on two lexical processing tasks (semantically related and semantically unrelated trials). Denser semantic connectivity (i.e., global level connectivity between near and far neighbors) positively associated with semantic interference during semantically-related lexical processing, while denser category structure (i.e., lower-level connectivity between near neighbors) facilitated lexical processing in semantically unrelated trials. In Experiment 2, a subset of the same children (N=49) returned at age 36 months and completed a comprehensive assessment of their language skills using the CELF-P2. Here, earlier measures of lexico-semantic connectivity and lexical processing best predicted age 3 language skill. The findings support accounts that early vocabulary structure and lexical processing skills promote continued growth in language.

Keywords

Vocabulary; lexico-semantic structure; lexical processing; longitudinal; eye-tracking

---

In the second half of the second year of life, children make substantial gains in their language abilities, including the ability to understand words in real-time (i.e. lexical processing), produce new words, and recognize semantic linkages between words. Notably, from 12 to 24 months, children typically progress from saying a handful of words to hundreds (Fenson, Dale, Reznick, Bates, & Thal, 1994). Vocabulary acquisition continues to accelerate such that most children start to produce many new words as they enter in the latter half of their second year (between 18 to 24 months of age; Goldfield & Reznick,

---

Author Note
I wish to extend my deepest gratitude to the many families who enthusiastically shared their child's language growth over 1.5 years. It is a tremendous privilege to learn from you. I also wish to thank the numerous research assistants and students who supported data collection, including: Katherine Chia, Christopher Cole, Megan Foley, Kathleen Hopkins, Ryan Peters, Edda Rodriguez, Angele Yazbec, and Courtney Yehnert.

1990). Across the same developmental window, toddlers increase their fluency in real-time word comprehension – a skill referred to here as "lexical processing" (Fernald et al., 1998). Together, these early vocabulary and lexical processing abilities forecast continued language growth, such that children who produce more words and interpret speech accurately and efficiently also have better language skills at preschool age and beyond (see Appendix A for a simplified schematic of relations among vocabulary size lexical processing and language outcomes; Friend et al., 2019; Marchman & Fernald, 2008; Peter et al., 2019; Smolak et al., 2021).

Further, as the *size* of the child's expressive lexicon changes, so does the semantic *structure* that characterizes it. By age two, children typically learn numerous word meanings that share semantic connections such that toddler's lexicons can be described as a dense network of word meanings (Hills et al., 2009; Peters & Borovsky, 2019). Relatively little is understood regarding how developmental changes in the structure of lexico-semantic networks are tied to changes in vocabulary, lexical processing, and later language outcomes. Illuminating these relations is critical to support a full account of how children rapidly acquire language and how early lexico-semantic skills support continued language growth.

The current project advances our understanding of these developmental dynamics in language acquisition through two experiments. Experiment 1 measures how lexical processing skills are influenced by vocabulary size and lexico-semantic structure at 18 and 24 months. Experiment 2 then assesses how these skills connect to broader developmental profiles by modeling whether lexical processing, vocabulary size and structure between 18 to 24 months forecasts language outcomes at age three (Experiment 2). The next sections first motivate Experiment 1 by illustrating how and why vocabulary structure and size relates to lexical processing skills. Then, the following section motivates Experiment 2 by exploring how these skills together (vocabulary size, structure and lexical processing) support longer-term outcomes in language skills.

Drivers of lexical processing

Lexical processing (i.e. real-time word recognition) involves a number of cognitive processes, including the ability to activate word meanings and select appropriate referents. Lexical processing improves across development, suggesting that children become more efficient at retrieving and identifying word meanings as they mature (Bergelson & Swingley, 2012; Fernald et al., 1998). In toddlers, lexical processing abilities are typically assessed in tasks measuring speed and or accuracy of a tactile response or gaze towards a labeled item in an array of objects (Fernald et al., 2008; Friend & Keplinger, 2003; Golinkoff et al., 1987). Numerous adaptions of these tasks have identified additional factors that influence speed and accuracy of lexical processing, such as the child's own vocabulary knowledge and the semantic relatedness of the distractor image(s) (Bergelson & Aslin, 2017; Borovsky et al., 2016a; Fernald et al., 2006). Below, three relevant factors that influence lexical processing are reviewed: (1) Semantic interference and facilitation, (2) Vocabulary size, and (3) Vocabulary structure. See also Appendix A for a schematic overview of the reviewed relations.

Semantic interference and facilitation

Lexical processing can be alternatively slowed or enhanced relative to other conditions. For example, lexical processing can be momentarily slowed by the presence of an unnamed distractor image that overlaps in taxonomic relation to a named target image (i.e., semantically related trials; Han & Li, 2020; Huettig & Altmann, 2005; Yee & Sedivy, 2006). This phenomenon, termed “semantic interference,” emerges in infancy (Bergelson & Aslin, 2017), and is thought to reflect the individual’s lexico-semantic vocabulary structure co-activates both the intended target meaning and related semantic neighbors, thereby creating “interference” that necessitates extra time to resolve and identify the appropriate target. Simply put, semantic interference emerges as a pattern of slower/less accurate performance in lexical processing tasks, and reflects simultaneous activation or competition from closely-related semantic neighbors in the comprehender’s vocabulary and from visually-present semantically-related objects. Conversely, semantic facilitation occurs in conditions that support more robust/efficient activation and retrieval of word meanings from the child’s semantic memory. Semantic facilitation appears in versions of lexical processing tasks where a semantic competitor is not in view (i.e., semantically unrelated trials), and where the child’s own vocabulary contains a relatively larger/denser semantic neighborhood surrounding a particular word meaning (described in Vocabulary structure section below).

Vocabulary size

Children’s lexical processing ability is also associated with vocabulary size. For example, lexical processing studies in toddlers and preschoolers repeatedly find that children who more accurately touch or view a named picture also have larger vocabularies (Fernald et al., 2006; Friend et al., 2019; Lany et al., 2018; Law & Edwards, 2015; Legacy et al., 2016, 2018; Mahr & Edwards, 2018; Marchman et al., 2016, 2010; Peter et al., 2019; Poulin-Dubois et al., 2013). Lexical processing skills also correlate with future vocabulary outcomes and suggest that that efficient processing and building a sizeable vocabulary are likely to support each other. Having a large vocabulary, for example, facilitates retrieval and recognition of known word meanings (thereby enhancing lexical processing skills). In turn, lexical processing efficiency frees up cognitive resources to support learning unknown words, thereby supporting vocabulary growth (Donnelly & Kidd, 2020; Friend et al., 2018, 2019; Lany, 2018; Marchman & Fernald, 2008; Peter et al., 2019).

Vocabulary structure

Further, lexical processing skills vary not only with the *size* of the child’s expressive vocabulary, but also its *structure*. Semantic priming studies are one way in which researchers have assessed how lexical processing varies with semantic structure. These types of studies support inferences about whether and to what degree semantic structure is represented in the early lexicon, by assessing how semantically related words impact lexical processing. For instance, the child’s processing of a spoken word can be slowed when viewing another semantically related object (Bergelson & Swingley, 2013), or enhanced when the spoken word is preceded by a related spoken word (e.g., Arias-Trejo & Plunket, 2013).

These semantic priming findings indicate that children represent their vocabulary among semantically related dimensions and that this structure, in turn, impacts early processing.

Importantly, the impact of these early structural representations on processing changes across 18 to 24 months. For example, in semantic priming studies, children between 18- to 21-months show variable performance, while 24-month-olds exhibit robust semantic priming (Arias-Trejo & Plunkett, 2013; Luche et al., 2014; Rämä et al., 2013; Torkildsen et al., 2006, 2007; Willits et al., 2013). Part of this early variability in semantic priming is additionally explained by vocabulary size, such that 18- to 20-month olds with relatively larger productive vocabularies exhibit more robust semantic priming effects compared to those with smaller vocabularies.

More direct evidence that semantic structure impacts early lexical processing comes from findings that the density of a semantic category surrounding a word facilitates lexical processing in 18- and 24-month-olds (Borovsky et al., 2016a; Borovsky & Peters, 2019). This semantic facilitation "boost" in lexical processing may be due to a variety factors, including more efficient and robust activation of target word meanings, coupled with appropriate inhibition of non-target meanings as the neighborhoods surrounding a word increase in density. For example, Chow and colleagues (2019) posit that the rapid increase in vocabulary around the time of the vocabulary spurt may be driven by (or result in) the need to develop inhibitory lexical activation mechanisms that prevent massive over-activation and interference in denser lexical neighborhoods that accompany vocabulary growth.

Thus, while there are promising clues that the structure of the child's vocabulary can influence lexical processing, there are many ways vocabulary structure could potentially impact processing. The next section introduces how vocabulary structure can be assessed at different levels of granularity, and this review motivates the need to assess vocabulary structure at various levels in the current project.

Levels of vocabulary structure.—Semantic structure in the lexicon exists at several levels (Steyvers & Tenenbaum, 2005), including at the word-level (i.e. words vary in the number of semantically-related neighbors), category-level (some categories have many known members, others are relatively sparse), and lexicon-level (some children have lexicons that are highly interconnected, other child lexicons may have fewer connections, in general, among words). These levels of semantic structure affect adult patterns of lexical processing. At the word level, directly connected semantic neighbors prime each other (i.e., *car* facilitates RTs to *truck* in priming tasks). Higher levels of semantic structure can affect processing across more distantly-connected lexico-semantic neighbors, as seen in indirect priming studies, where words like *cat* prime *cheese* via a shared neighbor, *mouse* (Hutchison, 2003; Neely, 1977). Lexical processing tasks involving lexical decision and categorization also exhibit semantic interference effects (with slower RTs) in response to words with many near neighbors, while semantic facilitation effects (faster RTs) emerge for words with more distantly related neighbors (Mirman & Magnuson, 2008).

Experiment 1 examines how semantic structure across these levels of granularity (word-level, category-level and lexicon-level) affect patterns of lexical facilitation and interference in 18-month-old and 24-month-old toddlers. The measurement approach and implications for psycholinguistic processing at each level of semantic granularity (category, word, and lexicon) are described below.

Category level structure.—Structure at the category level reflects density in taxonomically-related semantic neighbors, and may signal the child's interests (Ackermann et al., 2020). For example, some toddlers may have relatively dense category-level structure for ANIMALS, but not for VEHICLES, and vice versa. In the current project, category-level vocabulary structure is calculated as a relative measure of category density in six early-acquired taxonomic categories with items on the MBCDI, following procedures outlined in several other studies (Borovsky et al., 2016b, 2016a; Borovsky & Peters, 2019). First, the proportion of words in each category is rank-ordered from highest to lowest, and the top three categories are assigned to a "high" density condition, while the bottom three are assigned to a "low" density condition. Note that this is a relative (and not an absolute) measure of category density. The rationale for this relative measure is so that density can be scaled according to the child's own vocabulary size, as children with larger productive vocabularies (both within and across age) are likely to have overall higher absolute density measurements. Therefore, even children who produce relatively few words can contribute equally to higher and lower density measures, while controlling for vocabulary size. This category-level measure represents structure at a "medium" grain size, reflecting a taxonomic super-ordinate level of organization of near- and medium-distance neighbors and between lower-level word and higher-level lexicon structure measures, described below.

Graph-theoretic measures of word-level and lexicon-level structure.—In addition to category-level semantic structure, the current project also assesses word-level and lexicon-level structure. Both of these levels of semantic granularity are calculated using graph-theoretic approaches. Word-level structure reflects the number of semantic neighbors for any particular word, while lexicon-level structure captures the semantic connectivity across the child's entire lexico-semantic network. In this study, semantic relations between words are formalized as words that share semantic feature overlap. The following sections first motivate this semantic feature overlap approach, before returning to a detailed description of the word-level and lexicon-level metrics.

Identifying semantic neighbors by semantic features.: A number of studies have sought to model early vocabulary networks to describe early patterns in lexical acquisition in relation to productive vocabulary size. These studies have varied in their approach to establishing semantic connections among words, variously using: free association norms (Fourtassi et al., 2020; Hills et al., 2009); child-directed speech corpora derived co-occurrence statistics (Beckage et al., 2011); semantic features (Hills et al., 2009; Hills et al., 2009; Peters & Borovsky, 2019; Siew, 2019; Yurovsky et al., 2012); or a combination of all of these (Stella et al., 2017). All of these approaches yield appropriate measures of semantic structure in early lexicon, and studies that have sought to compare multiple semantic modeling methods in psycholinguistic studies have found a number of redundancies in among feature norm and distributional approaches (Riordan & Jones, 2011). Stella and colleagues (2017) modeled semantic structure in relation to lexical acquisition using both feature norm and association rating data, and found that both approaches significantly predicted acquisition in children under the age of two (i.e., the age range in which vocabulary networks are modeled in this study). Moreover, semantic-feature based approaches have a long-standing tradition within the psycholinguistic literature to model and understand cognitive phenomena associated with activation and representation of word meanings, dating back at least half a century (e.g. Collins & Loftus, 1975; Collins & Quillian, 1969). Semantic feature approaches have persisted due to their extensively replicated ability to explain behavioral and neural indices of performance in a range of psycholinguistic tasks (Chang et al., 2011; Grondin et al., 2009; McRae et al., 1997; Pexman et al., 2008; Rabovsky et al., 2021; Taylor et al., 2012). Compared to the adult literature, relatively less developmental work has focused on how lexical processing is impacted by featural overlap among word meanings, though there is some evidence that early lexical processing can be affected by perceptual (feature) overlap among word meanings (Arias-Trejo & Plunkett, 2010; Mani et al., 2013; Peters et al., 2021). For example, Arias-Trejo & Plunkett (2010) find toddlers' lexical processing is most impacted when a target object appears with a distractor that shares perceptual and taxonomic features in common, and not either one alone. Taken together, while there are reasons to suspect that featural representations support lexical processing even at earliest stages of language learning, there is a need to understand precisely when and how semantic-feature based models of early vocabulary structure explain patterns in lexical processing across development.

Therefore, the current study seeks to measure how processing is impacted by semantic structure in the child's own lexicon, by describing semantic connections as a function of semantic feature overlap using semantic feature norms (e.g., McRae et al., 2005) for all noun concepts in the MBCDI. In this feature approach, features are first selected based on normed rating of toddler accessibility, which range on a scale from 1 to 7, with 7 ranking as most highly accessible. Only features with a value of 4 or greater are then used to construct a lexico-semantic network for each participant using the individual nouns that the child is reported to produce (i.e. network "nodes"). Connections between words (i.e. network "links" or "edges") are created between all word pairs that share semantic feature overlap. These individual lexico-semantic networks are then used to describe word-level and lexicon-level measures of degree (DEG) and global clustering coefficient (GCC), respectively.

<u>**Word-level structure as degree.:**</u> To address the question of how the number of direct semantic neighbors for an individual word impacts its processing, the degree for each target word in the study was calculated as the number of words in the child's productive lexicon that share semantic overlap with the target word. Higher degree values indicate that a word has many semantic neighbors. From a graph-theoretic perspective, the degree of any individual node (i.e., word) is taken as an indicator of the potential for that node to communicate with many other nodes in the network. When this idea is translated to lexico-semantic networks, "communication" indicates the potential for a word to be influenced via spreading activation from a number of direct word neighbors. Word-level degree therefore represents the potential for interference or facilitation of lexical processing by near-neighbors that share direct semantic connections.

Lexicon-level structure as GCC.: To address the question of how the semantic connectivity of the child's overall lexicon impacts processing, lexicon-level structure was measured as global clustering coefficient, which represents the mean interconnectivity of words and their neighbors. GCC represents the probability that any randomly selected word pair out of the child's total vocabulary will be connected and therefore reflects the potential of both near and far neighbors to facilitate or interfere with lexical processing. A lexicon with a relatively higher GCC has more words and word-neighbors that are connected, and this structure is argued to support efficient word retrieval and recognition in adults and children (Griffiths et al., 2007; Steyvers & Tenenbaum, 2005)

(How) does lexical processing, vocabulary size and structure predict later language skills?—Building a sizeable and well-structured vocabulary is likely to support lexical processing skills, but how do these three lexico-semantic constructs (vocabulary size, structure and lexical processing) relate to longer term developmental profiles? One possibility is that building a large and well-structured vocabulary may influence lexical-processing abilities during this period of development but that these skills (processing, vocabulary size, and structure) fail to foster future language growth. Indeed, many aspects of the language input change as children age (Bornstein et al., 1999; Kavanaugh, 1982; Sokolov, 1993), which could necessitate a shift in skills that predict and promote later language growth. Alternatively, early lexico-semantic skills may promote later (and more complex) language abilities. This possibility is supported by evidence that the emergence of morpho-syntactic skills is rooted in building a sizeable lexicon (Bates & Goodman, 1997; Caselli et al., 1999; Marchman & Bates, 1994) and that individual differences in lexical processing skills relate to future language outcomes at preschool and school-age (Fernald & Marchman, 2012; Marchman & Fernald, 2008). This possibility is also supported by observations that parents adapt their speech complexity in response to growth in the child's skills, suggesting a direct connection between early language skills and later ones (e.g. Hani et al., 2013; Leung et al., 2021). To help clarify whether and how these skills measured at 18 and 24 months support longer-range language growth, Experiment 2 explores how lexical processing, vocabulary structure and size at 18 to 24 months predict longer-term preschool-age language ability at 36 months.

The current studies and potential outcomes—This project seeks to advance our understanding of the co-development of lexico-semantic processes and language development more broadly in several respects. Experiment 1 explores how the size and structure of the child's lexicon at 18 and 24 months supports their lexical processing skills. Then, Experiment 2 addresses whether and how these three measures (lexical processing, vocabulary size and vocabulary structure) predict later language outcomes at age three. To achieve these goals, a longitudinal sample of toddlers completed a set of tasks that assessed their early vocabulary size, structure, and lexical processing skills at 18 and 24 months (see Fig. 1: Experiment 1). Then, to relate these skills to language outcomes in later toddlerhood, participants returned at 36 months to complete a comprehensive assessment of language skills using the Level 1 subtests from Clinical Evaluation of Language Fundamentals – Preschool 2 (Semel et al., 2004; see Fig. 1: Experiment 2). The analyses are reported in two parts. First, Experiment 1 measures how performance on two lexical processing tasks (word recognition in semantically unrelated and semantically related contexts) is predicted by measures of vocabulary size and structure at 18 and 24 months. These analyses yield insights into how two mechanisms of lexico-semantic activation – facilitation and interference – are tied to developmental changes in the size and structure of toddlers' lexicons. If changes in vocabulary size (but not structure) explain unique variance in lexical processing tasks that tap into facilitation and interference, it would suggest that the mechanisms that support lexical retrieval and activation of word meanings mature in tandem with the child's skill in word learning/vocabulary building. Alternatively, it is possible that these analyses will reveal that vocabulary structure (but not size) influences performance on the lexical processing tasks. This outcome would suggest that the child's ability not simply to learn words, but to recognize and activate connections among those word meanings support lexical processing. With the outcomes of these analyses in hand, Experiment 2 then measures how earlier performance on all three constructs (lexical processing, vocabulary size and structure) connect to later language outcomes at age 3. Although it is expected that lexical processing skills and vocabulary size are likely to relate to later outcomes, it is unknown if vocabulary structure will explain additional unique variance in outcomes. If such a relation does appear, then it would suggest that building not simply a *sizeable* vocabulary but also a *well-connected* vocabulary mediates long-term growth in language skills.

Experiment 1

Methods

Participants—The project initially enrolled 18- to 20-month-old toddlers living near a small city to study developmental changes in vocabulary size, structure, and lexical processing and language outcomes at age 3. Enrolled families were invited to take part in the simple lexical processing tasks reported here at 18 and 24 months, and an outcome measure of language skills at 36 months. The results of the children in the lexical processing task at 18-months were previously reported in Borovsky and Peters (2019). 101 families with 18-month-olds completed the MBCDI and participated in the eye-tracking task, and 78 of these children were retained in the final analyses (23 were excluded for the following reasons: falling >1.5 SD below the mean on the cognitive subtest of the Bayley Scales of Infant Toddler Development (N=1); failure to meet language criteria (N=3); parent suspected an uncorrected hearing or vision concern (N=4); or child was receiving services for a motor, speech or language concern (N=5)). An additional 10 18-month-olds were also excluded due to fussiness or excessive data loss, defined as resulting in fewer than two trials in each experimental condition. Of these 78 children, 67 also returned at 24 months and completed the same lexical processing task, and their caregivers completed another MBCDI at that time. Of these 67 returning 24-month-olds, six were excluded from analysis due to fussiness or data loss (i.e., contributed fewer than two trials per condition for analysis), leaving a sample of 61 children who completed the experimental tasks at both 18 and 24 months. We initially report analyses on this dataset with respect to changes in processing and vocabulary size and structure in Experiment 1. In Experiment 2, we carry out further analyses on the subset of children (N=49) who returned to the lab again and completed a CELF outcome measure at 36 months. Demographics of the sample in Experiment 1 and Experiment 2 are reported in Table 1. A priori power analyses were carried out to determine the necessary sample size to measure a medium sized effect and interaction in the planned models for Experiments 1 and 2. These analyses indicated a need for minimum sample size of 55 to detect a medium size multiple regression effect (f^2=0.15), with alpha threshold at .05, and 80% power. While this sample threshold was achieved for Experiment 1 (N=61), sample attrition caused it to fall below this number in Experiment 2 (N=49). Therefore, a post-hoc sensitivity analyses was carried out to determine the achieved power for this Experiment 2. This analysis indicated that a sample size of 49 yields 80% power to detect a medium size multiple regression effect (f^2=0.168). Parents or guardians of children in the study provided written informed consent for participation and the Florida State University IRB approved the study “Development of Language Processing Skills” (#2017.21274).

Materials

Category and item selection: The materials for this task were identical to that reported in Borovsky and Peters (2019). As in that prior study, items were selected from six early-acquired semantic categories that contain a number of items that 18- to 24-month-olds commonly say, according to the Cross-Linguistic Lexical Norms Database (CLEX; Jørgensen, Dale, Bleses, & Fenson, 2010). The six category domain were: ANIMALS, CLOTHING, VEHICLES, BODY-PARTS, along with two food sub-categories: FRUIT and DRINKS. The MBCDI checklist includes the following number of items each of these subcategories respectively: 43, 28, 14, 27, 7, and 7. The experimental items included two items from each category, and were verified in CLEX to be produced by at least 70% of 18-month-olds. In unrelated trials, each object was paired with an item from an alternating category. In related trials, each object was paired with the other item in the same category. All participants saw trials with each object in related and unrelated pairings during the study. All items and item pairings are listed in Appendix B.

Auditory Stimuli: Spoken labels for the experimental task were recorded from a female speaker of Standard American English speaking in a child-directed voice, using a mono channel at 44.1 kHz. To ensure that all items were spoken at relatively similar rates and intensity, lexical items were normalized to the mean length of 1020 ms and to a mean intensity (70 dB). Experimental trials also began with a recording of the same speaker saying “Look!” which was modified to the same intensity (70 dB). The same speaker also provided recordings of other phrases that were used to support toddler’s interest during the study. These included tag phrases which were spoken after each experimental item, that included phrases like, “That’s cool!” “Great job!” as well as other encouraging phrases that appeared in between experimental trials like, “These pictures are fun!” “You’re doing great!” These encouraging phrases were all modified to the same mean intensity as the experimental items (70 dB).

Visual stimuli: Objects in the study were presented as 400 × 400 pixel photos on a 17” monitor with screen resolution of 1280 × 1024 pixels. Most images appeared on white backgrounds, except for several images that were not easily recognizable in isolation, such as an image of a “nose.” Images were selected to be representative of toddler’s visual experience, which we verified by consulting with parents and other laboratory members. For example, the image of a SHOE was selected to be a colorful toddler’s sneaker. See

Figure 2 for an illustration of visual stimuli in related and unrelated trials. During the study, toddlers also viewed other images to encourage ongoing attention to the task. These images included larger and smaller images of various cartoon characters that were accompanied by encouraging verbal stimuli.

Procedure

Offline assessments

Knowledge of experimental items.: The goal of the study was to explore how children understand words that are already familiar to them. Therefore, at each age, child knowledge of each experimental item was verified via parental report. Parents provided ratings of their child's knowledge of each experimental item on a scale of 1 (Definitely does not understand) to 4 (Definitely understands). Any items where the parent marked less than a '2' for the labeled target item were removed from analysis at that age. This criterion led to a removal of 90 trials at 18 months (4.7% of the dataset), and 6 experimental trials at 24 months (<0.13% of the dataset).

Cognitive assessment (18 months only).: This study sought to explore developmental changes and outcomes in children who show otherwise typical cognitive development. Therefore, at 18 months, all children completed the Cognitive subtest of the Bayley Scales of Infant and Toddler Development (Bayley-III; Bayley, 2005). Any child with a standard score <85 (1.5 SD below the mean SS) was excluded. One child was excluded for failing to reach this criterion. The standard scores for the remaining 61 children who completed the task at 18 and 24 months, ranged from 85–145 (M=113.1, SD=14.3)

Vocabulary assessment.—A central goal of the study was to explore how vocabulary size and structure influences lexical processing in related and unrelated contexts. Therefore, to support the calculation of these measures, parents completed a MacArthur-Bates Communicative Developmental Inventory, Words and Sentences (MBCDI; Fenson et al., 2007) form when their child was 18 and 24 months. This inventory is norm-referenced and standardized for children between 16 to 30 months of age and yields a vocabulary percentile score at each month of age within this range. Between 18 to 24 months of age, this instrument captures a relatively comprehensive snapshot of the child's expressive vocabulary (Mayor & Plunkett, 2011).

Percentile scores ranged from 1^{st}-94^{th} percentile (10–374 words produced) at 18 months ($M_{\%tile}$ = 44.3, $SD_{\%tile}$ = 23.8, M_{words}= 90.4, SD_{words}=77.45), and 6^{th}-99^{th} percentile (61–662 words produced) at 24 months ($M_{\%tile}$ = 52.4, $SD_{\%tile}$ = 24.5, M_{words}=302.5, SD_{words}=150). Figure 3a and 3d illustrate vocabulary percentile distribution (3a) and changes across age (3d). As in an earlier analysis of the 18-month-olds in this dataset, children were grouped into higher and lower vocabulary groups at each age based on a median split of productive vocabulary.

Experimental procedure: After an initial warm-up period in an adjacent room, the child was encouraged to sit in a car seat in front of a monitor on a moveable arm. The monitor was placed in front of a curtain, and an eye-tracking camera was affixed directly below the monitor. The caregiver and experimenter each sat on either side of the child as they watched the experimental images together. The caregiver was instructed not to name the images on the screen. A second experimenter remained out of view behind the curtain to monitor the focus and calibration of the eye-tracking apparatus and experiment.

Before the experiment, the eye tracker was focused and calibrated using a 5-point routine. To verify that the calibration remained stable throughout the study, each experimental trial was initiated with a centrally-presented "drift-check" image (30 × 30 pixels in size). Once the child fixated on the drift check, the experimental trial began with images of the target and distractor object on the left and right side of the screen. The goal of this silent preview was to allow the child to become familiar with the object images and their locations. After 2000 ms of silent image preview, a 100 × 100 pixel gaze-contingent central stimulus appeared, paired with a verbal stimulus, "Look!" The target object label was spoken only after the child fixated towards the central image. This gaze-contingent procedure served several goals. First, it ensured that experimental trials only proceeded when the child was attentive to the verbal and visual stimuli on the screen. Second, it allowed every trial to start from when the child was fixating towards the same central location, thereby increasing confidence that sustained eye-movements towards the target stimulus were mediated by the auditory stimulus and thereby mitigating a potential source of experimental noise (see Delle Luche, Durrant, Poltrock, & Floccia, 2015 for an excellent demonstration of this point). Figure 4 illustrates the sequence of events in a trial.

Approach to analyses: The analytic goals were to illuminate how vocabulary size and structure influenced processing of semantically unrelated and semantically related trials across age. Therefore, children who contributed data at both points only were included to facilitate understanding of these changing patterns. To maintain consistency within the literature, the analytic approach in this paper follows the statistical practices in the earlier report of the 18-month-olds only. The approach to calculating measures of task performance, vocabulary size and structure are described in detail below.

Vocabulary size.: As prior semantic priming studies at these ages (reviewed above) suggest that higher and lower vocabulary groups display different patterns in priming, participants were assigned to higher and lower vocabulary groups. At each age, toddlers were assigned to either a High (N_{18mos} = 29, N_{24mos} = 32) or Low productive vocabulary group (N_{18mos} = 32, N_{24mos} = 29) according to a median split of MBCDI productive vocabulary score at that age. While vocabulary percentile was included as a continuous measure in full models, analysis were additionally carried out for related and unrelated trials for high and low vocabulary median split groups separately. The MBCDI data was also used to calculate three measures of lexico-semantic structure in each toddlers' productive vocabulary following procedures set out in prior studies and described below.

Category-level structure (Category Density).: The goal of this measure is to highlight relative differences in the semantic density of items in the surrounding semantic category neighborhood while controlling for each child's overall vocabulary size. Category density was calculated as follows: (1) Calculate the proportion of words the child produces in each of the six categories from which the experimental items were selected; (2) Rank order the category proportions from highest to lowest proportion; (3) Assign the top three proportion categories to a HIGH density condition, and the bottom three to a LOW density condition; and (4) Resolve any #3 and #4 rank ties by assigning both to the HIGH condition. Table 2 reports the distribution of HIGH and LOW density categories at 18 and 24 months.

Word-level structure (Degree).: Degree reflects the number of directly connected semantic neighbors for each lexical item that appeared in the study. Word degree was calculated for each child as the number of other nouns that they were reported to produce on the MBCDI that shared at least two semantic features in common with each lexical item that was presented in the experiment. Semantic features were derived from sets of semantic feature norms (e.g., McRae et al., 2005). To ensure that these measures of semantic structure emerge from features that are accessible to young children, features were only included in the models if they were rated to be accessible to toddlers (mean rating of 4 or higher on a scale from 1 to 7, with 7 being very highly accessible). Word degree ranged from 0 to 42. Concretely, a word degree of zero for an experimental item would indicate that the child does not produce any words that share semantic feature overlap for that word, according to parental report using the MBCDI. A higher degree value, such as 10, would indicate that the child produces 10 other words that share semantic feature overlap. As children age, so does the mean degree for all experimental items, illustrated in Figure 3b and 3e.

Lexicon-level structure (Global clustering coefficient).: Global clustering coefficient (GCC), is a graph-theoretic index of the overall semantic connectivity for each child's productive vocabulary. GCC is calculated by dividing the number of "closed" connected word triplets by the "open and closed" triplets of connected words in each child's productive lexicon. A word triplet is defined as any set of three words that share some kind of feature overlap, and an "open" triplet is one where 2 out of the 3 words are fully connected (i.e., cat-dog, cat-mouse is an open triplet, with cat as the "hub"). A triplet is defined as "closed" if all three words are directly connected (i.e. share semantic feature overlap) to each other (e.g. car-truck, car-helicopter, truck-helicopter). GCC values range between 0 and 1, with 0 indicating that there are no closed triplets in the child's vocabulary, and 1 indicating that every triplet is closed. Importantly, it is possible to have a vocabulary structure where a child has many pairwise connections between words (having a high word degree), but a low GCC, indicating that the neighbors of individual words are not connected. Therefore, GCC represents a more global measure of overall connectivity between both near and far neighbors of individual words, whereas Degree is a local measure of a word's direct neighbors. Figure 3c and 3f illustrates distribution of GCC values in participants at 18 and 24 months.

Eye-movement data analysis.: The analytic plan followed precedent established in a prior paper exploring vocabulary size and structure dynamics on lexical processing on the 18-month-olds in this sample Borovsky & Peters, 2019. As outlined in this prior paper, it was hypothesized that children would experience facilitation as a function of vocabulary size and structure in semantically unrelated conditions, whereas these same factors might promote semantic interference in the semantically related conditions. Given these differences in task demands and dynamics, performance on related and unrelated trials was explored separately.

Also, as prior studies had noted group differences between children with relatively larger and smaller vocabularies, as defined by median split, children at each age were divided into higher and lower vocabulary size for separate analyses for each vocabulary group. Vocabulary percentile was included as a continuous variable in full analyses exploring the impact of vocabulary size across all participants. To maintain consistency and facilitate comparisons, these analytic approaches are carried forward in the current analysis, which now also includes age as an additional factor in the statistical models.

For each condition, the time course of looking toward the target and distractor is plotted to illustrate the timing and magnitude of processing patterns in the relevant conditions across age. It is important to note that all fixation plots start at 0 indicating that toddlers were viewing neither the target nor the distractor at the onset of the verbal label. This pattern is due to the gaze-contingent trial design, where the target label was spoken only after the child initially fixated towards a central stimulus. To quantify how looking patterns vary across size and structure, accuracy of looking is calculated using a standard time window from 300 to 2,000 ms post-label onset. Note that this time window was used in prior analysis of the 18-month-olds, and is a typical time window that is used for lexical processing analyses in children of this age (Fernald et al., 2008). Accuracy is defined as a log-adjusted Target preference (i.e. log-gaze proportion ratio), which is calculated as the log-proportion of total fixation to the Target divided by Distractor [log (P_Target / P_Distractor)] during the defined 300–2000 ms time window. Because log values are undefined at 0, all 0 values in the numerator and denominator were replaced with a small value (.001). Positive log-gaze ratios indicate a target preference (and successful word recognition), while negative values indicate a distractor preference. The log-gaze adjustment has been adopted by a number of eye-tracking researchers as this transformation reduces statistical problems with homogeneity of variance and linear independence compared to raw fixation proportions that vary between 0 and 1 (Arai, van Gompel, & Scheepers, 2007; Knoeferle & Kreysa, 2012; Wienholz & Lieberman, 2019). This study was not pre-registered, and analysis code is available at osf.io/kt9gs.

Results

Time course visualization—The time course of lexical processing in relation to the three semantic structure measures at 18 and 24 months is illustrated in Figure 5a. As expected, these plots reveal that children at both ages rapidly and accurately identified the appropriate target object, typically within the first 500–1000 ms of the onset of the spoken label. Fixations towards the target exceeded those towards the distractor by the end of analysis period (2000 ms) in all conditions. Importantly, the timing of these gaze dynamics illustrate that the pre-defined standard analytic time window of 300 to 2000 ms post label onset encompasses the relevant phenomena (see Fernald et al, 2008 for more details on this time window recommendation). Additionally, inspection of the time course also reveals that looking patterns potentially vary by age and vocabulary structure measure. The following statistical analyses seek to quantify these illustrated patterns.

Unrelated trial recognition accuracy—These analyses sought to measure how lexical processing patterns were facilitated as a function of vocabulary structure and skill across age. Relevant patterns as a function of vocabulary size and structure are illustrated in Figure 6. Performance on each trial was measured as log-gaze trial accuracy across a time window from 300 to 2,000 ms post label onset (as described in 'Approach to Analysis' above). Additionally, categorical measures were deviation coded (High = 0.5, Low=−0.5; Barr, 2019; Barr, 2008), and continuous variables were standardized. These measures were then entered into a linear-mixed-effects regression (LMER) model using the *lme4* library in R. To facilitate consistency in the literature, these analyses follow the statistical modeling approach in Borovsky & Peters (2019), which initially included random effects of Participants and Items (i.e., words in the study). However, as in this prior paper, not all of these models converged, but all models converged when including Items as a random effect. Results of the model are listed in Table 2, and the model formula is:

$$LogGaze \sim (GCC + Degree + Category\ Density) * Vocabulary * Age\ Group + (1 |\ Items)$$

The statistical model illustrates several important effects. First, replicating prior findings, the positive intercept (marginally significant) indicates that the sample as a whole successfully recognized the words in the task, suggesting this task was appropriate across the ages tested. A positive effect of Category Density illustrated that lexical processing was facilitated in categories with higher density, also consistent with other reviewed studies with children in this age range. Unexpectedly, there was not an effect of Age on recognition accuracy, indicating that children did not significantly change in their overall performance on this lexical processing task between 18 to 24 months. No other main effects or interactions reached significance in the full model with all participants.

Unrelated trial accuracy by vocabulary group.—Following the analytic plan, based on prior publication of this task with 18-month-olds only (Borovsky & Peters, 2019), and prior semantic priming findings which suggested that priming effects vary substantially across higher and lower vocabulary groups, participants were further divided into higher and lower vocabulary groups to explore how semantic structure influenced word recognition accuracy in each group across age. In the higher vocabulary group (Table 3b), a significant positive intercept term indicated robust recognition of the target items. However, unlike in the full group analyses, category density did not significantly facilitate lexical processing for the higher vocabulary group. On the other hand, in the lower vocabulary group, there was not a significant, positive intercept effect, suggesting that this group, as a whole, exhibited less consistency in target word recognition, though lexical processing was facilitated for items in denser semantic categories (indicated by a significant category density term in the model in Table 3c). This lower vocabulary group also exhibited a marginal interaction of category density and age ($p = .068$), which is driven by a change in the impact of density on lexical recognition in unrelated trials across age, with a clear density effect at 18 months, but not 24 months.

Related trial recognition accuracy—In contrast to the semantically unrelated task where it was hypothesized that semantic structure should facilitate lexical processing, it was expected that semantic structure should interfere with lexical processing of a known object in the presence of a semantically related distractor. Task performance as a function of semantic structure and vocabulary skill is illustrated in Figures 5b, and 7. The analyses were carried out using the identical statistical approach and formula as in the unrelated trials, and model results are reported in Table 4.

As in the semantically unrelated task, the significant positive intercept indicated that children successfully fixated towards the labeled target object. There was also a marginal negative effect of Age Group on the task, indicating that children experienced greater semantic interference at 24 months compared to 18 months. There was a main effect of Category Density on performance, with a negative estimate indicating that High density items to experienced more semantic interference (reduced accuracy), compared to Low density items. Category Density and GCC also interacted with Age. Follow-up Tukey analyses of the Category density x Age interaction effect suggest it was driven by reduced interference (i.e., higher accuracy) for lower-density items at 18 months, compared to 24 months, $\beta(SE) = 0.71(0.26)$, $t = 2.78$, $p_{tukey\text{-}adj} = .028$. Two further models explored how lexical processing in semantically-related trials was impacted by age and semantic structure on Higher and Lower vocabulary groups separately. These model results are illustrated in Table 4b and 4c, and reveal that the main effect of Category density, and interactions of Age with Category Density and GCC were driven by the higher vocabulary group. In the higher vocabulary group, there was a significant main effect of Age group, with negative estimates indicating the higher vocabulary group experienced greater semantic interference (reduced accuracy) in the task as they got older. Secondly, there was an interaction of Category Density x Age Group which post-hoc Tukey analyses suggested was driven by a tendency for low density items to elicit greater interference at 24 vs. 18 months, $\beta(SE) = 1.16(0.34)$, $t = 3.38$, $p_{tukey\text{-}adj} = .004$. A further interaction of GCC across age was driven by changes in the influence of GCC on lexical processing at each age, and this effect was particularly strong for children in the higher vocabulary group. As illustrated in Figures 5b and 7, 24-month-olds with relatively more (vs. less) interconnected productive lexicons (i.e., higher GCC) experienced reduced interference on recognizing words in this task, and this effect was particularly driven by the higher vocabulary group.

Discussion

This experiment was designed to test how mechanisms of semantic interference and facilitation play out during lexical processing in trials where there either is or is not a semantically related distractor item. The findings illustrate that, as expected, children appropriately recognized the appropriate target items across both tasks, though the dynamics of this process varied across task, and according to the size and structure of productive vocabulary.

In unrelated trials, denser category structure facilitated lexical processing, replicating prior research on semantically unrelated trials (Borovsky et al., 2016a; Borovsky & Peters, 2019). This effect was stronger among children with smaller vocabularies, suggesting that, the facilitative effects of category-level density in the lexicon may be particularly powerful when children have relatively smaller vocabulary sizes for their age. It is important to note that category density is a relative measure, such that higher and lower density categories are relative to the child's own vocabulary size, suggesting that the number of items that a child needs to produce to have a category count as a “sparse” or low-density category for a child who produces many words might be the threshold for a “dense” or high-density category for a child who produces many words. Yet, proportionally speaking, denser categories have a particularly stronger influence among children who know fewer words, suggesting that interventions that focus on teaching words within a category might be especially beneficial for children with early delays in expressive vocabulary skills. The relative measure of density has a number of advantages, including that it is not correlated with vocabulary size, as it is tied to relative differences in the child’s own vocabulary networks, thereby allowing all children to contribute to both higher and lower density categories at each age (see Borovsky et al., 2016 for expanded discussion).

In related trials, GCC and category density also affected performance, replicating and extending prior findings. Most notably, children with more interconnected vocabularies (higher GCC), experienced a reduction in semantic interference, and this effect was particularly strong for the higher vocabulary group. However, it should be noted that the range of GCC at 24 months in the higher vocabulary group is much more limited compared to children at different ages and with smaller productive vocabularies. Nevertheless, the overall relation between GCC and processing at 24 months (reduced interference with higher connectivity) does exist across the group as a whole (see Figure 7) as evident in the statistically significant Age x GCC interaction in the full sample analysis as well. The GCC measure, unlike the other measures of lexico-semantic structure, reflects structure from both near and distant neighbors, whereas the other structure measures reflect more local semantic neighborhood connections. This influence of distant neighbors on GCC suggests that as children’s vocabularies grow with age they experience reduced semantic interference from far neighbors.

While these results highlight important connections among lexical processing and developing structure as the lexicon builds over toddlerhood, they do not illustrate whether and how these skills might impact developmental processes over a longer developmental period. The next study pursues a logical next step in addressing whether and how the lexico-semantic skills that were measured in Experiment 1 supported continued language growth into age three, with a subsequent longitudinal follow up at 36 months on a comprehensive assessment of language skills that goes beyond simple vocabulary size, using a norm-standardized assessment of language.

Experiment 2

All participants in Experiment 1 were invited to return to the laboratory to participate in a comprehensive standardized assessment of language skills at 36 months. Here, it was expected that measures of lexical processing in unrelated contexts, as in prior work, would predict later outcomes. However, prior studies have yielded inconsistent findings regarding whether measures of expressive vocabulary size in early toddlerhood reliably predict later language outcomes, particularly at the level needed to address clinical outcomes (Thal et al., 2013). Therefore, while it was expected that vocabulary size should show positive relations with age three language outcomes, it was unclear whether vocabulary size would remain a primary indicator after accounting for unique variance contributed by lexical processing and structure. Finally, as prior work (reviewed above) had suggested that children with early expressive language delays show poorer metrics of vocabulary structure, it was expected that GCC would be positively associated with outcomes at 18 and 24 months, such that children with more interconnected lexicons would show higher age 3 language outcome scores.

Methods

Participants—Of the 61 children who were included in the analysis in Experiment 1, 49 returned at 36 months of age and completed an outcome assessment of language skill using the Clinical Evaluation of Language Fundamental, Preschool 2 (CELF-P2).

Assessments

CELF-P2 CLS: Children completed all Level 1 subtests of the Clinical Evaluation of Language Fundamentals – Preschool 2 (CELF-P2; Wiig, Secord & Semel, 2006) which is a norm-referenced standardized assessment of language skills that is suitable for children at this age. Each child's performance on the task was used to derive a norm-referenced standard core language score (CLS), which has a mean standard score (SS) of 100 and a standard deviation of 15. These standard scores indicate where a child falls along the normal distribution of language skills at this age. The CLS is derived by calculating the child's combined performance on several subtests that assess various aspects of age-appropriate language abilities. The subtests in the CLS assessment include: sentence structure, word structure and expressive vocabulary. Sentence structure items involve selecting an image in response to a spoken sentence, with item difficulty increasing with sentence length and complexity. For example, an early item on the test might involve viewing an array of four pictures, and the child is asked to point to an item that shows "it smells good" which only requires an ability to interpret a single element of the sentence, whereas a more difficult item would involve interpreting sentences with more complex phrasal structure that describe more complex events, like "point to *She drank the milk before she ate the sandwich,"* in an array of pictures involving a girl in various combinations of consuming a meal. Word structure items evaluate the child's ability to use morpho-syntax in various inflected, derivational, comparative and pronominal forms by asking the child to complete a verbal sentence completion (cloze) task. For example, a child might complete a sentence while viewing an image of a boy who is standing and another who is sitting, while the experimenter points to each image and says, "This boy is standing. This boy is …." The child would receive points for producing the correct progressive –ing form. Expressive vocabulary items evaluate the child's expressive vocabulary by eliciting a verbal label for illustrations. Therefore, together, the subtests provide a comprehensive evaluation of the child's ability to comprehend and use age-appropriate grammatical and semantic forms of language. Further, this test is often used by clinicians who seek to assess whether a preschool-aged child qualifies for a diagnosis of a language disorder, with a standard score of 85 or lower indicating a likely language delay at this age. In the analyses, the CELF P2 CLS is used as a continuous dependent measure of language outcome, and in a secondary exploratory analysis, it is used a categorical measure of presence/absence of a language delay (with a threshold of 85 SS, representing a group-wise distinction between children who fall below 1 SD of age-standard language skill).

MBCDI: The analyses included the same MBCDI-derived vocabulary size and GCC values at 18 and 24 months as described in Experiment 1. We do not include MBCDI-derived measures of category density and word degree, as these describe semantic structure that surround individual items on the experimental task, and thus do not describe overall individual differences in the size or structure of the child's lexicon, as indexed by GCC and vocabulary percentile measures.

Lexical processing performance: The mean log-gaze accuracy measures in semantically related and unrelated trials for each subject at 18 and 24 months from Experiment 1 were also included as predictors of language outcome. This study was not pre-registered, and analysis code is available at osf.io/kt9gs.

Results

The analytic goals were to explore whether and how language skill at 36 months is predicted by earlier measures of vocabulary size, structure, and lexical processing. First, to illustrate key relationships among the variables, the pairwise correlations and data density distributions across all measures are reported in Figure 8. Several notable patterns are apparent in these correlation tables. As expected, vocabulary percentile at 18 months is strongly correlated with vocabulary at 24 months. GCC was positively correlated with concurrent vocabulary size at 18 months, indicating that children with larger vocabularies also had greater semantic connectivity in their lexicons at this age. Part of the relation between GCC at 18 months appears to be driven by a group of children with smaller vocabularies who also tended to have GCC values around 0, representing a lack of semantic clustering at small vocabulary sizes for this group. However, clustering was not entirely absent in all children with smaller vocabularies, as illustrated on the correlation plot. Importantly, GCC at 18 months correlates with future language measures, including MBCDI at 24 months, and CELF-P2 at 36 months. At 24 months, GCC and MBCDI are not correlated, though GCC at this age does predict CELF-P2 outcome at 36 months – though, the direction of this relation is reversed, as compared to the relation between GCC at 18 months and CELF-P2 scores. At the same time, MBCDI percentile measures at 18 and 24 months does not correlate with later CELF-P2 language outcome measure at 36 months.

The next analyses sought to address whether and how the vocabulary-derived and lexical processing measures predicted unique variance in 36 month outcomes, while controlling for other factors. A multiple linear regression model was constructed to predict continuous 36 month language outcomes scores (CELF-P2 CLS standard score) as a function of vocabulary size and structure as well as the semantically related and unrelated lexical processing measures at 18 and 24 months (Table 5a, Model 1). Overall, the model predicted significant variance in 36-month language outcomes, $R^2_{multiple}$=.28, $F(8,46)$=2.21, p = .043. Individual predictors of GCC at 18 and 24 months predicted later CELF-P2 language scores at 36 months, but vocabulary percentiles did not. Interestingly, GCC at 18 months was positively associated with 36 month language outcome, such that 18 month-olds with more densely connected vocabularies had better age 3 language skills. At 24 months, the directionality of this relation changed: here, children who had more densely connected vocabularies had poorer age 3 language outcomes. Performance on the semantically-unrelated lexical processing task at 18 months also forecasted 36 month language scores such that 18-month-olds who more accurately recognized known words in unrelated contexts had better 36 month outcomes.

Finally, as performance on the CELF P2 is typically used by clinicians to identify children with language delays, the next analyses sought to explore whether the same predictors of vocabulary size, structure and processing from the previous models of *continuous* language outcomes, might also serve utility in predicting *categorical* low-language outcomes. First, 36 month-olds were categorized into either a "No Delay" status (having a SS >=85 on the CELF-P2 CLS, N=48), or with a "Language Delay" status (LD; with a SS<85 on the CELF-P2 CLS, N=7). Figure 9 illustrates differences between children with and without LD on all measures, where the most pronounced group differences appear as a function of GCC at 18 months. A binomial logistic regression was then conducted to predict language delay status as a function of vocabulary percentile, GCC, and performance in semantically-related and unrelated lexical processing tasks at 18 months. This analysis, reported in Table 5b (Model 2), indicated that GCC at 18 months and performance on the semantically unrelated trials at 18 months significantly predicted LD status. Specifically, children with lower GCC scores, and those who exhibited poorer performance on lexical processing in semantically unrelated trials, (see Figure 9) were more likely to fall within the LD group.

General discussion

This project comprises two longitudinal experiments which build on each other. The first (Experiment 1) enhances our understanding of how developmental changes in the size and structure of toddlers' vocabularies impact the mechanisms that support lexical processing, and the second (Experiment 2) explores how these skills together (lexical processing, vocabulary size, and structure) support ongoing growth of language skills into preschool. This work seeks to fill a gap between what have been traditionally two distinct areas of psycholinguistic inquiry. The first area has focused on key determinants and developmental trajectories of childhood language outcomes, and the other has delineated the dynamics and mechanisms that drive lexical activation, primarily in adults. The first domain has pointed to promising early interconnections between vocabulary size and lexical processing in supporting the continued growth of language skills into preschool and school-age, while the second has suggested that lexical recognition is supported not only by building a sizeable vocabulary, but as a function of the network of meaning surrounding a word. While a number of studies have focused on how semantic structure supports multiple measures of lexical processing in adulthood, relatively less is understood regarding how processing and vocabulary skills grow in tandem with vocabulary structure at the outset of learning. Like adults, vocabulary structure in the toddler's developing lexicon exists across a range of granularities, which can be measured at a local level, with the number of semantic neighbors for an individual item (word degree), building into taxonomically-oriented semantic categories (category density) and scaling up to connectivity that exists between immediate and distant neighbors in the entire lexicon (GCC). While studies of adult lexical processing had indicated that measures that account for semantic distance / granularity differentially affect processing in mature lexicons with many thousands of words, relatively little was known about how these factors affect lexical processing at the earliest stages of lexical learning. Therefore, Experiment 1 focused on how lexical processing was affected by vocabulary structure at several levels of granularity in the child's lexicon. The second experiment then sought to build on the sizeable literature which now suggests that lexical processing and vocabulary size are both important predictors for continued growth in the child's skills, while exploring how a new dimension – that of semantic structure – also contributes to longer-range language growth. The results across both studies provide useful constraints on potential mechanisms that support sustained language growth during this important period of development.

Effects of vocabulary structure on lexical processing—Vocabulary structure was measured at three levels: at the word-level, category-level, and total lexicon-level. Here, several findings are notable. First, the lack of word-level structure effects across the board suggests lexical processing mechanisms might not be influenced by direct (or nearest) semantic neighbors insomuch as it is driven by activation in a broader network of near and far semantic neighbors at the category and lexicon-level that are not necessarily directly connected to the target word. However, this lack of effect should be interpreted cautiously, and a great deal more work is needed to systematically explore and verify this conclusion for a lack of directly connected near-neighbor influence here.

The category and lexicon-level effects on lexical processing are also noteworthy, and support several inferences about the mechanisms that underlie early lexical activation across development. The category-level effects on facilitation and interference, particularly in the younger age group, suggest that categorical knowledge may be a powerful mechanism for learning and understanding words before age 2, which may, in turn, help to further promote vocabulary skills. For example, other work highlights that cues to basic-level category membership such as shape, can support early word learning (Perry et al., 2016; Smith et al., 2002). The current research adds to that work, by suggesting that super-ordinate category information may promote word learning by supporting lexical processing. Future work in this area is needed to identify whether these category density effects emerge from factors intrinsic to the child, such as their own interest (Ackermann et al., 2019) or from extrinsic factors, such as linguistic structure in child directed speech (Huebner & Willits, 2018), or in some mixture of the two.

There were also notable developmental shifts in the relation between lexicon-level structure and semantic interference during lexical processing. At 24 months, higher lexicon-level connectivity was associated with reduced interference, while the opposite was true at 18 months. At both ages, these effects were most strongly driven by children with relatively larger productive vocabularies. Prior work had reported that adult lexical processing is facilitated by distant semantic neighbors, while near semantic neighbors inhibit lexical processing (Mirman & Magnuson, 2009). The current findings provide additional constraint on the emergence and development of these mechanisms. At 18 months, near lexical neighbors (as indexed by semantic category density) exert a particularly powerful effect on lexical processing. Here, co-activation of near neighbors (i.e., other words in the same semantic category) appear to increase interference (by reducing accuracy in related trials) and boosting lexical facilitation (by increasing accuracy in unrelated trials). However, as the child's lexicon grows, distant lexical connections that would be indexed by GCC also start to also impact lexical processing, resulting in reduced interference for children with more highly interconnected lexicons (higher GCC) at 24 months. This developmental reduction in interference across age and connectivity seems consistent with accounts of the emergence of inhibitory links as vocabulary size grows (Chow et al., 2019; Mayor & Plunkett, 2014). As posited by Chow and colleagues (2019), the rapid increase in vocabulary around the time of the vocabulary spurt, may be driven by (or result in), the need to develop inhibitory lexical activation mechanisms that prevent massive over-activation and interference from the greater number of potential near and far neighborhood connections that result from acquiring a larger lexicon. Thus, the current findings support the idea that changes in the size and structure of the child's lexicon are tied to the emergence of mechanisms of semantic interference over this period. Further work is necessary to elucidate how these mechanisms of facilitation and interference shift and stabilize with continued growth in lexico-semantic knowledge throughout childhood and into adulthood.

Connecting lexical processing, vocabulary size and structure to later language skills—The final goal of this project was to explore whether and how toddlers' vocabulary size, vocabulary structure, and lexical processing predicted age 3 language skills. While it was expected that there would be a significant, positive correlation between earlier productive vocabulary skills and age 3 language outcome, none of the analyses identified this statistical connection. This lack of effect may be partially explained by variability that comprises typical vocabulary growth over the course of the second year (Fenson et al., 1994). Indeed, many children who show early delays in expressive vocabulary tend to 'catch up' over time, and expressive vocabulary size on its own in toddlerhood has been a somewhat inconsistent marker of later language delays at preschool and school-age (Rescorla & Dale, 2013; Thal et al., 2013). Overall, when considering vocabulary size in toddlerhood, the current findings point to importance of incorporating other measures in early assessment, such as vocabulary structure and lexical processing measures.

There were several novel relations between vocabulary structure, lexical processing and age 3 language outcomes. Lexical connectivity (GCC) at 18 and 24 months predicted continuous and categorical measures of language skill at age 3, though the direction of this connection changed across this period. At the youngest age tested (18 months), toddlers with more densely inter-connected vocabulary (as indexed by GCC) also had better age 3 language outcomes, whereas the direction of this relation had swapped by the time children turned two. This developmental change is driven by a dramatic shift in lexical connectivity among the toddlers with the most extreme GCC values at 18 months. Here, as illustrated in Figure 3f, while most children maintained relatively stable GCC values ranging between .4 - .6 across age, several children who started with the lowest GCC at 18 months then then converged to GCC values within this range by 24 months. This same group who progressed from the lowest GCC values at 18 months, and high GCC values at 24 months had the poorest CELF outcomes at age 3 (illustrated in Figure 8). One concern is that this convergence may be driven by a potential reduction in variability in the composition of vocabulary as children learn more words and hit ceiling on the MBCDI. However, closer inspection of the vocabulary data suggested this was not the case. Children exhibited a range of percentiles (and corresponding vocabulary sizes) at both ages, as illustrated in Figure

3a and 3d. Correspondingly, raw vocabulary also ranged considerably range across age ($Range_{words}$ = 10–374, M_{words} = 90.4 at 18 months; $Range_{words}$ = 61–662, M_{words}= 302.5 at 24 months), and, importantly, none of the children reached ceiling (680 words) on this instrument, thereby leaving room for individual variability in vocabulary composition on this assessment.

Rather, it appears that GCC values represent a kind of “goldilocks” marker for continued language growth into preschool. GCC values that converge around .5 at 18 months (and not too much in either direction) portend potential for positive language skill outcomes by preschool age. Too much or too little connectivity may signify potential for delay. Alternatively, given that children with more extreme GCC values at 18 months also exhibited relatively larger shifts in GCC by 24 months (Figure 3f), it is possible that dramatic change in lexico-semantic connectivity during this period could signal risk for poorer outcomes. This idea may also explain the age-reversed pattern between GCC and age 3 language skill. Future work is needed with larger longitudinal samples to distinguish whether risk for poorer language outcomes is connected to early “extremism” in lexico-semantic network structure or developmental instability in structure over time, as either possibility could be consistent with the current data. Further, recalling GCC’s connection to a reduction in semantic interference at 24 months in Experiment 1, GCC may also contribute to long term growth by spurring the development of inhibitory links in the lexicon, which may help to support sustained growth in language skills over time. This connection suggests that a possible fruitful direction for further inquiry may come from studies that explore, in detail, the drivers of lexical inhibition and connectivity in the lexicon.

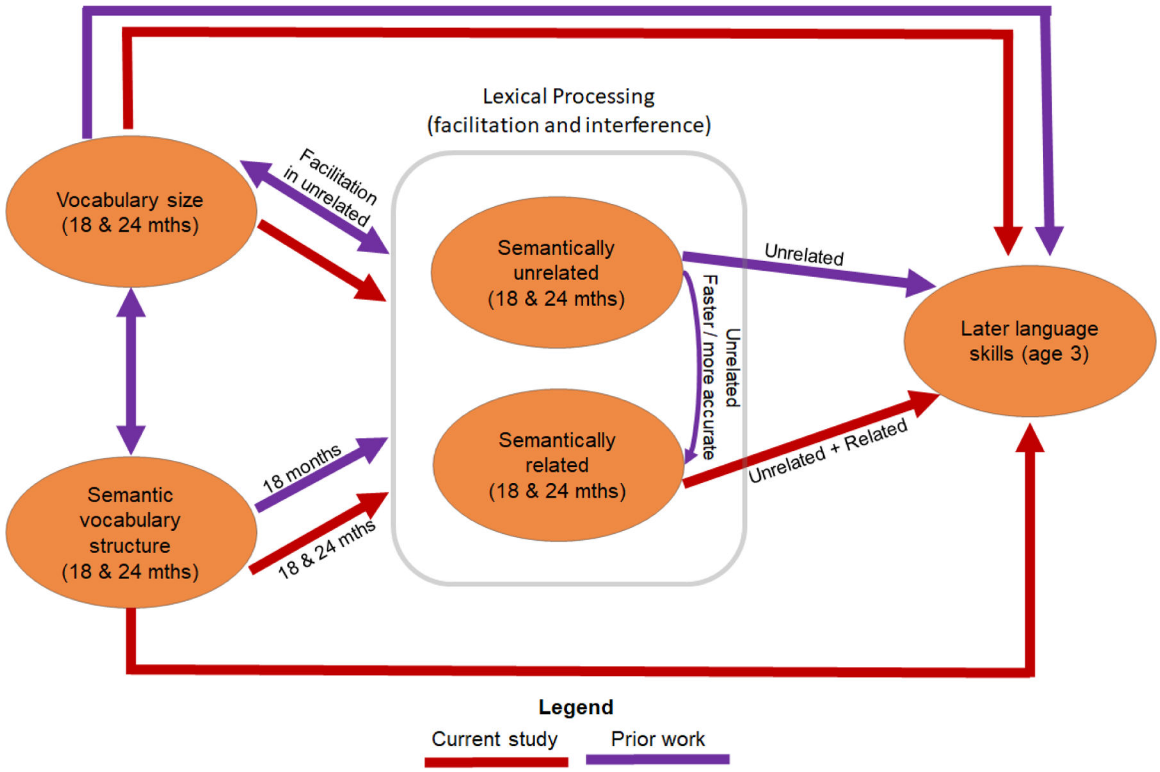
Measures of lexical processing also predicted later outcomes. Simple pairwise correlations indicated that accuracy in lexical processing in semantically related trials at 18 months significantly predicted outcomes at age 3. However, after controlling for vocabulary size and structure, only lexical processing in semantically unrelated trials predicted outcome measures. These findings replicate and extend prior findings which suggest that lexical processing skills predict later language skills, suggesting that the child’s ability to efficiently activate and recognize a named word is a critical component in language learning, and potentially supports not only continued vocabulary growth, but other, more complex language skills as well (Fernald & Marchman, 2012; Lany et al., 2018; Marchman & Fernald, 2008; but cf. Donnelly & Kidd, 2020).

In sum, this work represents an initial attempt to disentangle how several lexico-semantic skills that change dramatically over the second year contribute to continued patterns in language growth. The emerging findings suggest that mechanisms that underlie lexical processing associate with concomitant changes in vocabulary, and that these skills have important consequences for continued language skills at age 3. Importantly, the findings add to a growing body of evidence that point to the promise of supplementing typical measures of vocabulary size with those that also probe developments in vocabulary structure over time (Borovsky et al., 2021), as both types of measures are likely to interact in important ways that illuminate the mechanisms and knowledge that contribute to the vast individual differences in early language growth.

Acknowledgments

This research was supported by National Institutes of Health grants: DC013638, DC018593, and HD052120. This study was not pre-registered. Analysis code can viewed on osf.io/kt9gs.

Appendix A.



Schematic of constructs referred to in the introduction and relation to prior work and the current study.

Appendix B.

All target items and their distractor pairings in related and unrelated trials.

Target Item Category	Target Item	Related Distractor	Unrelated Distractor
ANIMALS	Bird	Dog	Nose
	Dog	Bird	Teeth
BODY PARTS	Teeth	Nose	Dog
	Nose	Teeth	Bird
CLOTHING	Diaper	Shoe	Airplane
	Shoe	Diaper	Car
DRINK	Juice	Milk	Banana
	Milk	Juice	Apple

Target Item Category	Target Item	Related Distractor	Unrelated Distractor
FRUIT	Apple	Banana	Milk
	Banana	Apple	Juice
VEHICLES	Airplane	Car	Diaper
	Car	Airplane	Shoe

References

Ackermann L, Hepach R, & Mani N (2020). Children learn words easier when they are interested in the category to which the word belongs. Developmental Science, 23(3). 10.1111/desc.12915

Arai M, van Gompel RPG, & Scheepers C (2007). Priming ditransitive structures in comprehension. Cognitive Psychology, 54(3), 218–250. 10.1016/j.cogpsych.2006.07.001 [PubMed: 16973144]

Arias-Trejo N, & Plunkett K (2010). The effects of perceptual similarity and category membership on early word-referent identification. Journal of Experimental Child Psychology, 105(1–2), 63–80. 10.1016/j.jecp.2009.10.002 [PubMed: 19931091]

Arias-Trejo N, & Plunkett K (2013). What's in a link: Associative and taxonomic priming effects in the infant lexicon. Cognition, 128(2), 214–227. 10.1016/j.cognition.2013.03.008 [PubMed: 23688648]

Barr DJ (2019). Coding categorical predictor variables in factorial designs. Supporting Materials and Blog Posts for Barr (2008) https://talklab.psy.gla.ac.uk/tvw

Barr DJ (2008). Analyzing 'visual world'eyetracking data using multilevel logistic regression. Journal Of Memory And Language, 59(4), 457–474. 10.1016/j.jml.2007.09.002

Bates E, & Goodman JC (1997). On the Inseparability of Grammar and the Lexicon: Evidence from Acquisition, Aphasia and Real-Time Processing. Language and Cognitive Processes, v12 n5–6 p.

Bayley N (2005). Bayley Scales of Infant and Toddler Development - Third Edition. Harcourt Assessment.

Beckage N, Smith L, & Hills T (2011). Small worlds and semantic network growth in typical and late talkers. PLoS ONE, 6(5), e19348. 10.1371/journal.pone.0019348 [PubMed: 21589924]

Bergelson E, & Aslin RN (2017). Nature and origins of the lexicon in 6-mo-olds. Proceedings of the National Academy of Sciences of the United States of America, 114(49), 12916–12921. 10.1073/pnas.1712966114 [PubMed: 29158399]

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.

Bornstein MH, Tamis-LeMonda CS, & Haynes OM (1999). First words in the second year: Continuity, stability, and models of concurrent and predictive correspondence in vocabulary and verbal responsiveness across age and context. Infant Behavior and Development, 22(1), 65–85. 10.1016/S0163-6383(99)80006-X

Borovsky A (2021). Analysis of Lexico-semantic structure in vocabulary and its links to lexical processing in toddlerhood and language outcomes at age three Retrieved from osf.io/kt9gs/?view_only=a31739c708ae49d2af2d8d9cec86edd0

Borovsky A, Ellis EM, Evans JL, & Elman JL (2016a). Semantic structure in vocabulary knowledge interacts with lexical and sentence processing in infancy. Child Development, 87(6), 1893–1908. 10.1111/cdev.12554 [PubMed: 27302575]

Borovsky A, Ellis EM, Evans JL, & Elman JLJL (2016b). Lexical leverage: Category knowledge boosts real-time novel word recognition in two-year-olds. Developmental Science, 19(6), 918–932. 10.1111/desc.12343 [PubMed: 26452444]

Borovsky A, & Peters RE (2019). Vocabulary size and structure affects real-time lexical recognition in 18-month-olds. PLOS ONE, 14(7), e0219290. 10.1371/journal.pone.0219290 [PubMed: 31295282]

Borovsky A, Thal D, & Leonard LB (2021). Moving towards accurate and early prediction of language delay with network science and machine learning approaches. Scientific Reports, 11, 8136–8147. 10.1038/s41598-021-85982-0 [PubMed: 33854086]

Caselli C, Casadio P, & Bates E (1999). A Comparison of the Transition from First Words to Grammar in English and Italian. Journal of Child Language, v26 n1 p69.

Chang K. min K., Mitchell T, & Just MA (2011). Quantitative modeling of the neural representation of objects: How semantic feature norms can account for fMRI activation. NeuroImage, 56(2), 716–727. 10.1016/j.neuroimage.2010.04.271 [PubMed: 20451625]

Chow J, Aimola Davies AM, Fuentes LJ, & Plunkett K (2019). The vocabulary spurt predicts the emergence of backward semantic inhibition in 18-month-old toddlers. Developmental Science, 22(2). 10.1111/desc.12754

Collins AM, & Loftus EF (1975). A spreading-activation theory of semantic processing. Psychological Review, 82(6), 407–428.

Collins AM, & Quillian MR (1969). Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior, 8(2), 240–247. 10.1016/S0022-5371(69)80069-1

Delle Luche C, Durrant S, Poltrock S, & Floccia C (2015). A methodological investigation of the Intermodal Preferential Looking paradigm: Methods of analyses, picture selection and data rejection criteria. Infant Behavior and Development, 40, 151–172. 10.1016/j.infbeh.2015.05.005 [PubMed: 26176183]

Donnelly S, & Kidd E (2020). Individual differences in lexical processing efficiency and vocabulary in toddlers: A longitudinal investigation. Journal of Experimental Child Psychology, 192, 104781. 10.1016/j.jecp.2019.104781 [PubMed: 31981753]

Ellis EM (2014). Lexical and cognitive processing in early language delay. Dissertation Abstracts International: Section B: The Sciences and Engineering, 75(4-B(E)), No-Specified.

Fenson L, Dale PS, Reznick JS, Bates E, & Thal D (1994). Variability in early communicative development. . Monographs of the Society for Research in Child Development, 59(Serial no, 242).

Fenson L, Marchman VA, Thal DJ, Dale PS, Reznick JS, & Bates E (2007). MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual, Second Edition. Brookes Publishing.

Fernald A, Perfors A, & Marchman VA (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. Developmental Psychology, 42(1), 98–116. 10.1037/0012-1649.42.1.98 [PubMed: 16420121]

Fernald A, Zangl R, Portillo AL, & Marchman VA (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In Sekerina I, Fernandez E, & Clahsen H (Eds.), Language Processing in Children Benjamins.

Fernald Anne, & Marchman VA (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. 10.1111/j.1467-8624.2011.01692.x [PubMed: 22172209]

Fernald Anne, Pinto JP, Swingley D, Weinberg A, & McRoberts GW (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. Psychological Science, 9(3), 228–231. 10.1111/1467-9280.00044

Fourtassi A, Bian Y, & Frank MC (2020). The Growth of Children's Semantic and Phonological Networks: Insight From 10 Languages. Cognitive Science, 44(7), e12847. 10.1111/cogs.12847 [PubMed: 32621305]

Friend M, & Keplinger M (2003). An infant-based assessment of early lexicon acquisition. Behavior Research Methods, Instruments, & Computers : A Journal of the Psychonomic Society, Inc, 35(2), 302–309. 10.3758/bf03202556

Friend M, Smolak E, Liu Y, Poulin-Dubois D, & Zesiger P (2018). A cross-language study of decontextualized vocabulary comprehension in toddlerhood and kindergarten readiness. Developmental Psychology, 54(7), 1317–1333. 10.1037/dev0000514 [PubMed: 29620386]

Friend M, Smolak E, Patrucco-Nanchen T, Poulin-Dubois D, & Zesiger P (2019). Language status at age 3: Group and individual prediction from vocabulary comprehension in the second year. Developmental Psychology, 55(1), 9–22. 10.1037/dev0000617 [PubMed: 30359059]

Goldfield BA, & Reznick JS (1990). Early lexical acquisition: rate, content, and the vocabulary spurt. J Child Lang, 17(1), 171–183. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=2312640 [PubMed: 2312640]

Golinkoff RM, Hirsh-Pasek K, Cauley KM, & Gordon L (1987). The eyes have it: lexical and syntactic comprehension in a new paradigm*. In J. Child Lang (Vol. 14).

Griffiths TL, Steyvers M, & Firl A (2007). Google and the Mind. Psychological Science, 18(12), 1069–1076. 10.1111/j.1467-9280.2007.02027.x [PubMed: 18031414]

Grondin R, Lupker SJ, & McRae K (2009). Shared features dominate semantic richness effects for concrete concepts. Journal of Memory and Language, 60(1), 1–19. 10.1016/j.jml.2008.09.001 [PubMed: 20046224]

Han H, & Li X (2020). Degree of conceptual overlap affects eye movements in visual world paradigm. Language, Cognition and Neuroscience, 1–9. 10.1080/23273798.2020.1797838

Hani HB, Gonzalez-Barrero AM, & Nadig AS (2013). Children's referential understanding of novel words and parent labeling behaviors: Similarities across children with and without autism spectrum disorders. Journal of Child Language, 40(5), 971–1002. 10.1017/S0305000912000426 [PubMed: 23021075]

Hills TT, Maouene M, Maouene J, Sheya A, & Smith LB (2009). Longitudinal analysis of early semantic networks: Preferential attachment or preferential acquisition? Psychological Science, 20(6), 729–739. 10.1111/j.1467-9280.2009.02365.x [PubMed: 19470123]

Hills TT, Maouene M, Maouene J, Sheya A, & Smith L (2009). Categorical structure among shared features in networks of early-learned nouns. Cognition, 112(3), 381–396. 10.1016/j.cognition.2009.06.002 [PubMed: 19576579]

Huettig F, & Altmann G (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. Cognition, 96(1), B23–B32. [PubMed: 15833303]

Hutchison KA (2003). Is semantic pniming due to association strength or feature overlap? A microanalytic review. Psychonomic Bulletin & Review, 10(4), 785–813. 10.3758/BF03196544 [PubMed: 15000531]

Jørgensen RN, Dale PS, Bleses D, & Fenson L (2010). CLEX: A cross-linguistic lexical norms database*. Journal of Child Language, 37(2), 419–428. 10.1017/S0305000909009544 [PubMed: 19570318]

Kavanaugh RD |Jirkovsky AM (1982). Parental Speech to Young Children: A Longitudinal Analysis. Merrill-Palmer Quarterly, 28(2), 297–311.

Knoeferle P, & Kreysa H (2012). Can speaker gaze modulate syntactic structuring and thematic role assignment during spoken sentence comprehension? Frontiers in Psychology, 3. 10.3389/fpsyg.2012.00538

Lany J (2018). Lexical-processing efficiency leverages novel word learning in infants and toddlers. Developmental Science, 21(3), e12569. 10.1111/desc.12569 [PubMed: 28597549]

Lany J, Giglio M, & Oswald M (2018). Infants' Lexical Processing Efficiency is Related to Vocabulary Size by One Year of Age. Infancy, 23(3), 342–366. 10.1111/infa.12228

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

Legacy J, Zesiger P, Friend M, & Poulin-Dubois D (2016). Vocabulary size, translation equivalents, and efficiency in word recognition in very young bilinguals. Journal of Child Language, 43(4), 760–783. 10.1017/S0305000915000252 [PubMed: 26044885]

Legacy J, Zesiger P, Friend M, & Poulin-Dubois D (2018). Vocabulary size and speed of word recognition in very young French-English bilinguals: A longitudinal study. Bilingualism (Cambridge, England), 21(1), 137–149. 10.1017/S1366728916000833 [PubMed: 29416429]

Leung A, Tunkel A, & Yurovsky D (2020). Parents fine-tune their speech to children's vocabulary knowledge 10.31234/osf.io/prz8g

Luche CD, Durrant S, Floccia C, & Plunkett K (2014). Implicit meaning in 18-month-old toddlers. Developmental Science, n/a-n/a. 10.1111/desc.12164

Mahr T, & Edwards J (2018). Using language input and lexical processing to predict vocabulary size. Developmental Science, 21(6), e12685. 10.1111/desc.12685 [PubMed: 29781230]

Mani N, Johnson E, McQueen JM, & Huettig F (2013). How yellow is your banana? toddlers' language-mediated visual search in referent-present tasks. Developmental Psychology, 49(6), 1036–1044. 10.1037/a0029382 [PubMed: 22845828]

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

Marchman VA, Adams KA, Loi EC, Fernald A, & Feldman HM (2016). Early language processing efficiency predicts later receptive vocabulary outcomes in children born preterm. Child Neuropsychology, 22(6), 649–665. 10.1080/09297049.2015.1038987 [PubMed: 26031342]

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

Marchman VA, Fernald A, & Hurtado N (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish-English bilinguals. Journal of Child Language, 37(4), 817–840. 10.1017/S0305000909990055 [PubMed: 19726000]

Mayor J, & Plunkett K (2011). A statistical estimate of infant and toddler vocabulary size from CDI analysis. Developmental Science, 14(4), 769–785. 10.1111/j.1467-7687.2010.01024.x [PubMed: 21676097]

Mayor J, & Plunkett K (2014). Infant word recognition: Insights from TRACE simulations. Journal Of Memory And Language, 71(1), 89–123. 10.1016/j.jml.2013.09.009 [PubMed: 24493907]

McRae K, deSa VR, & Seidenberg MS (1997). On the nature and scope of featural representations of word meaning. Journal of Experimental Psychology-General, 126(2), 99–130. [PubMed: 9163932]

McRae K, Cree GS, Seidenberg MS, & Mcnorgan C (2005). Semantic feature production norms for a large set of living and nonliving things. Behavior Research Methods, 37(4), 547–559. 10.3758/BF03192726 [PubMed: 16629288]

Mirman D, & Magnuson JS (2008). Attractor dynamics and semantic neighborhood density: Processing is slowed by near neighbors and speeded by distant neighbors. Journal of Experimental Psychology: Learning, Memory, and Cognition, 34(1), 65–79. 10.1037/0278-7393.34.1.65 [PubMed: 18194055]

Mirman D, & Magnuson JS (2009). Dynamics of activation of semantically similar concepts during spoken word recognition. Memory & Cognition, 37(7), 1026–1039. 10.3758/MC.37.7.1026 [PubMed: 19744941]

Neely JH (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. Journal of Experimental Psychology: General, 106(3), 226–254.

Perry LK, Axelsson EL, & Horst JS (2016). Learning What to Remember: Vocabulary Knowledge and Children's Memory for Object Names and Features. Infant and Child Development, 25(4), 247–258. 10.1002/icd.1933

Peter MS, Durrant S, Jessop A, Bidgood A, Pine JM, & Rowland CF (2019). Does speed of processing or vocabulary size predict later language growth in toddlers? Cognitive Psychology, 115, 101238. 10.1016/j.cogpsych.2019.101238 [PubMed: 31539813]

Peters R, & Borovsky A (2019). Modeling early lexico-semantic network development: Perceptual features matter most. Journal of Experimental Psychology: General, 148(4), 763–782. 10.1037/xge0000596 [PubMed: 30973265]

Peters RE, Kueser JB, & Borovsky A (2021). Perceptual Connectivity Influences Toddlers' Attention to Known Objects and Subsequent Label Processing. Brain Sciences, 11(2), 163. 10.3390/brainsci11020163 [PubMed: 33513707]

Pexman PM, Hargreaves IS, Siakaluk PD, Bodner GE, & Pope J (2008). There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. Psychonomic Bulletin and Review, 15(1), 161–167. 10.3758/PBR.15.1.161 [PubMed: 18605497]

Poulin-Dubois D, Bialystok E, Blaye A, Polonia A, & Yott J (2013). Lexical access and vocabulary development in very young bilinguals. International Journal of Bilingualism, 17(1), 57–70. 10.1177/1367006911431198

Rabovsky M, Schad DJ, & Abdel Rahman R (2021). Semantic richness and density effects on language production: Electrophysiological and behavioral evidence. Journal of Experimental Psychology. Learning, Memory, and Cognition, 47(3), 508–517. 10.1037/xlm0000940 [PubMed: 33090840]

Rämä P, Sirri L, & Serres J (2013). Development of lexical–semantic language system: N400 priming effect for spoken words in 18- and 24-month old children. Brain And Language, 125(1), 1–10. 10.1016/j.bandl.2013.01.009 [PubMed: 23435193]

Rescorla LA, & Dale PS (2013). Late Talkers: Language development, interventions, and outcomes Paul H. Brookes.

Riordan B, & Jones MN (2011). Redundancy in perceptual and linguistic experience: Comparing feature-based and distributional models of semantic representation. Topics in Cognitive Science, 3(2), 303–345. 10.1111/j.1756-8765.2010.01111.x [PubMed: 25164298]

Semel EM, Wiig EH, & Secord W (2004). Clinical Evaluation of Language Fundamentals-Preschool, Second Edition. Pearson.

Siew CSQ (2019). Feature distinctiveness effects in language acquisition and lexical processing: Insights from megastudies. Cognitive Processing, 1, 3. 10.1007/s10339-019-00947-6

Smith LB, Jones SS, Landau B, Gershkoff-Stowe L, & Samuelson L (2002). Object name learning provides on-the-job training for attention. Psychological Science, 13(1), 13–19. 10.1111/1467-9280.00403 [PubMed: 11892773]

Smolak E, Hendrickson K, Zesiger P, Poulin-Dubois D, & Friend M (2021). Visual and haptic responses as measures of word comprehension and speed of processing in toddlers: Relative predictive utility. Journal of Experimental Child Psychology, 203, 105032. 10.1016/j.jecp.2020.105032 [PubMed: 33221662]

Sokolov JL (1993). A Local Contingency Analysis of the Fine-Tuning Hypothesis. Developmental Psychology, 29(6), 1008–1023. 10.1037/0012-1649.29.6.1008

Stella M, Beckage NM, & Brede M (2017). Multiplex lexical networks reveal patterns in early word acquisition in children. Scientific Reports, 7, 46730. 10.1038/srep46730 [PubMed: 28436476]

Steyvers M, & Tenenbaum JB (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. Cognitive Science, 29(1), 41–78. 10.1207/s15516709cog2901_3 [PubMed: 21702767]

Taylor KI, Devereux BJ, Acres K, Randall B, & Tyler LK (2012). Contrasting effects of feature-based statistics on the categorisation and basic-level identification of visual objects. Cognition, 122(3), 363–374. 10.1016/j.cognition.2011.11.001 [PubMed: 22137770]

Thal DJ, Marchman VA, & Tomblin JB (2013). Late-talking toddlers: Characterization and prediction of continued delay. In Rescorla LA & Dale PS (Eds.), Late talkers: Language Development, Interventions, and Outcomes (pp. 169–202). Paul H. Brookes.

Torkildsen J. von K., Sannerud T, Syversen G, Thormodsen R, Simonsen HG, Moen I, Smith L, & Lindgren M (2006). Semantic organization of basic-level words in 20-month-olds: An ERP study. Journal of Neurolinguistics, 19(6), 431–454. http://www.sciencedirect.com/science/article/B6VDV-4JCCJV6-1/2/9adf0bd96bc934335821f69b3f540176

Torkildsen J, Syversen G, Simonsen HG, Moen I, & Lindgren M (2007). Electrophysiological correlates of auditory semantic priming in 24-month-olds. Journal of Neurolinguistics, 20(4), 332–351. http://www.sciencedirect.com/science/article/B6VDV-4NJ26N3-1/2/4b62002593b0ef5a756a45746bed0b18

Wienholz A, & Lieberman AM (2019). Semantic processing of adjectives and nouns in American Sign Language: effects of reference ambiguity and word order across development. Journal of Cultural Cognitive Science, 1–18.

Willits JA, Wojcik EH, Seidenberg MS, & Saffran JR (2013). Toddlers activate lexical semantic knowledge in the absence of visual referents: evidence from auditory priming. Infancy, 18(6), 1053–1075.

Yee E, & Sedivy JC (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition; Journal of Experimental Psychology: Learning, Memory, and Cognition, 32(1), 1–14. [PubMed: 16478336]

Yurovsky D, Bion RAH, Smith LB, & Fernald A (2012). Mutual exclusivity and vocabulary structure. In Miyake N, Peeble D, & Cooper RP (Eds.), Proceedings of the 34th annual meeting of the Cognitive Science Society (pp. 1197–1202). Cognitive Science Society.

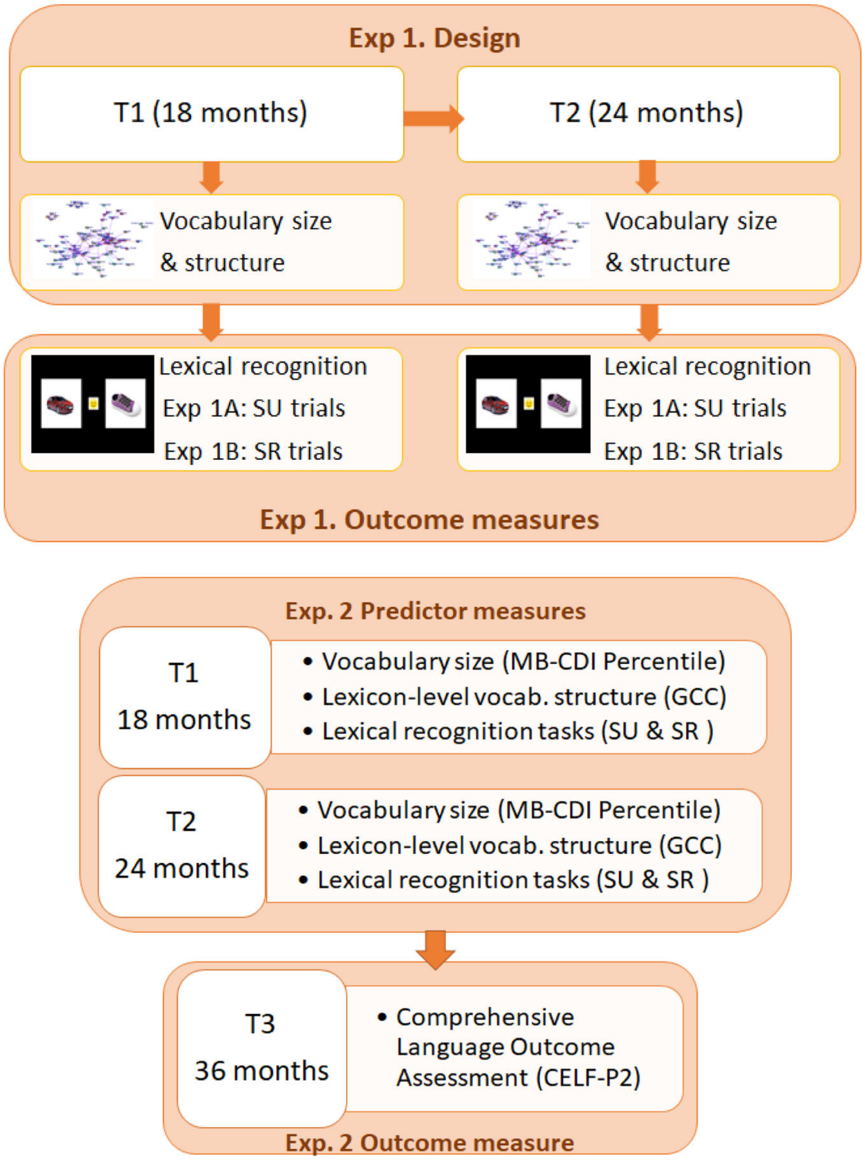


Figure 1.
Overview of design and measures in Experiments 1 and 2. Car image courtesy of © User:Thesupermat / Wikimedia Commons / CC-BY-SA-3.0 (background removed). Shoe image edited from https://www.maxpixel.net/Fashion-Sneakers-Cute-Sports-Shoes-Childrens-Shoes-3057511/CC01.0.

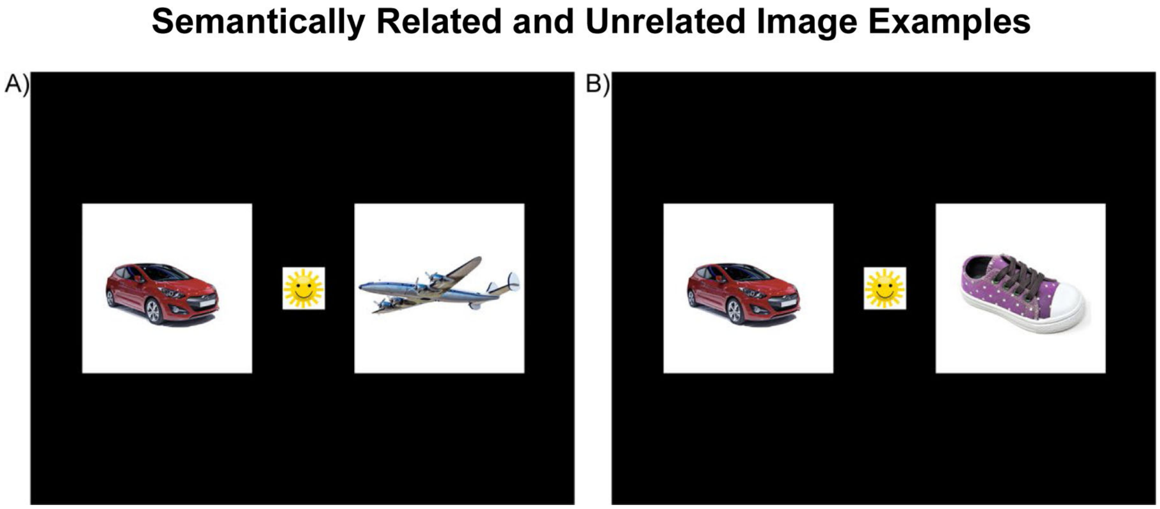


Figure 2. Examples of visual stimuli in each condition.

(A) Illustrates semantically related trials, where Target and Distractor images were semantically related. (B) Depicts semantically unrelated trials, where Target and Distractor did not share category membership. Due to licensing restrictions, similar, but not identical images of depicted objects appeared in the study. Car image courtesy of © User:Thesupermat / Wikimedia Commons / CC-BY-SA-3.0 (background removed). Remaining illustrated images were edited from images selected from https://www.maxpixel.net/ / CC0 1.0. This image is adapted under a CC BY license from “Vocabulary size and structure affects real-time lexical recognition in 18-month-olds” by Borovsky, A. and Peters, R.E., PLOS One, 14(7),p. 7, https://doi.org/10.1371/journal.pone.0219290.g001

Figure 3.

Upper plots illustrate distribution of (A) Vocabulary Percentile, (B) Degree, and (C) GCC at each age, and lower plots show change in (D) Vocabulary Percentile, (E) Degree and (F) GCC for each individual subject across age. Color in panels A–C denote participant age, while color in panels C and D indicate the starting value of Vocabulary Percentile (D) Degree (E) and GCC (F) at 18 months for each child to illustrate patterns of developmental change across participants in the data.

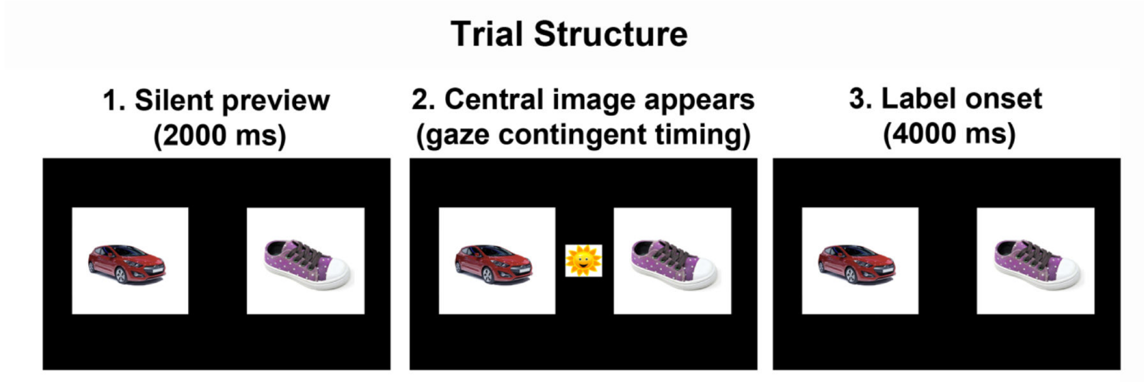


Figure 4. Illustration of sequence of events in a trial.

Trial structure is illustrated in 1–3. Before the onset of the trial, a central image appears to direct the child's attention to the screen. Then (1) Trial begins with silent preview of two images. (2) A central stimulus appears, and remains until child fixates on it. (3) Central image disappears and the label of the target object is spoken *(Car! That's cool!).* Images remain on the screen for 4000 ms post label onset. Car image courtesy of © User:Thesupermat / Wikimedia Commons / CC-BY-SA-3.0 (background removed). Shoe image edited from https://www.maxpixel.net/Fashion-Sneakers-Cute-Sports-Shoes-Childrens-Shoes-3057511/CC01.0.

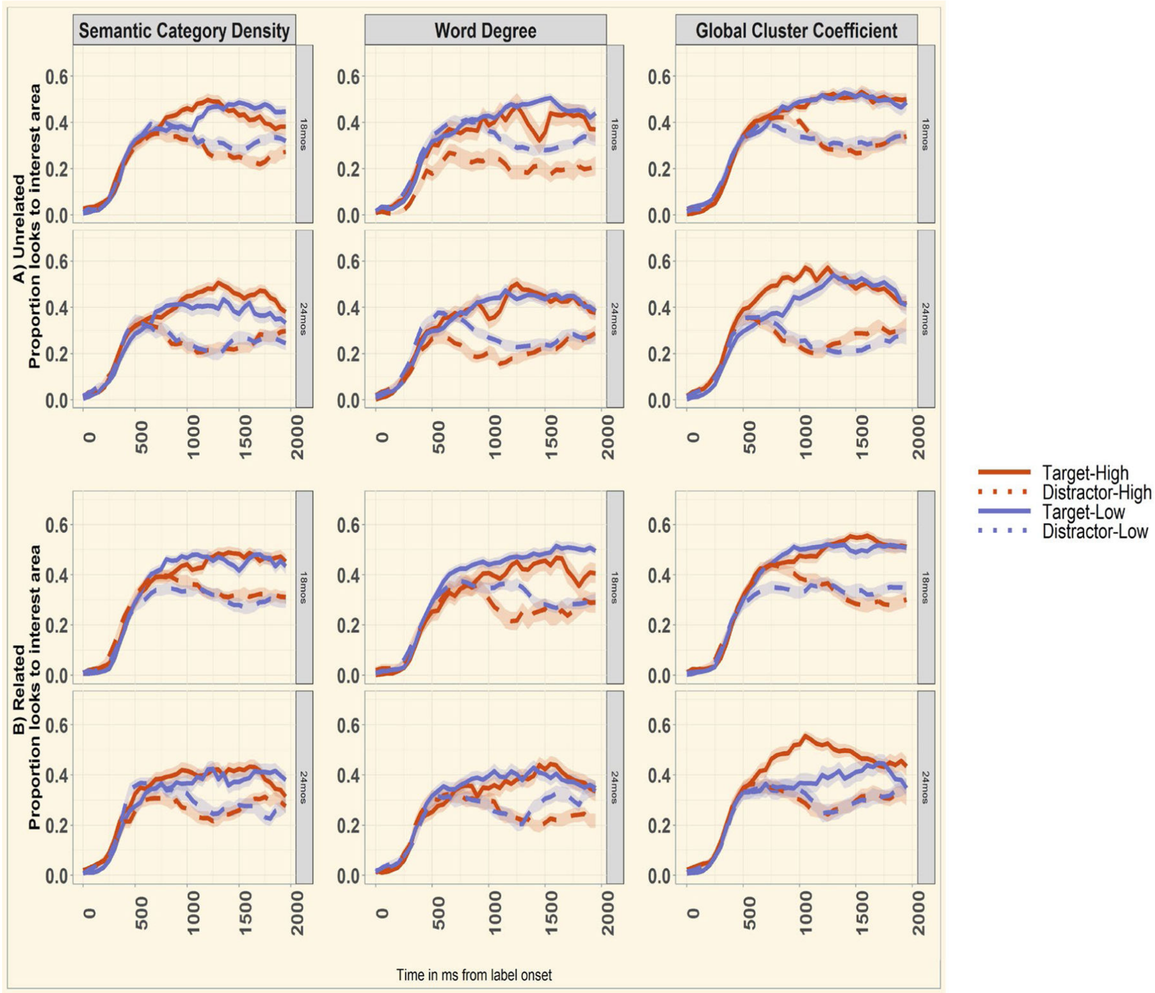


Figure 5.

Time course plots illustrating patterns of recognition as a function of semantic structure at 18 and 24 months. Plots depict proportion of looks towards the Target or Distractor object in 50 ms intervals starting from label onset (0 ms) to 2000 ms post label onset. High/Low for category density groups indicate looks to words which belong to high or low density categories. High/Low for word degree and GCC represent a median split of by item (for degree) and by child/lexicon (for GCC).

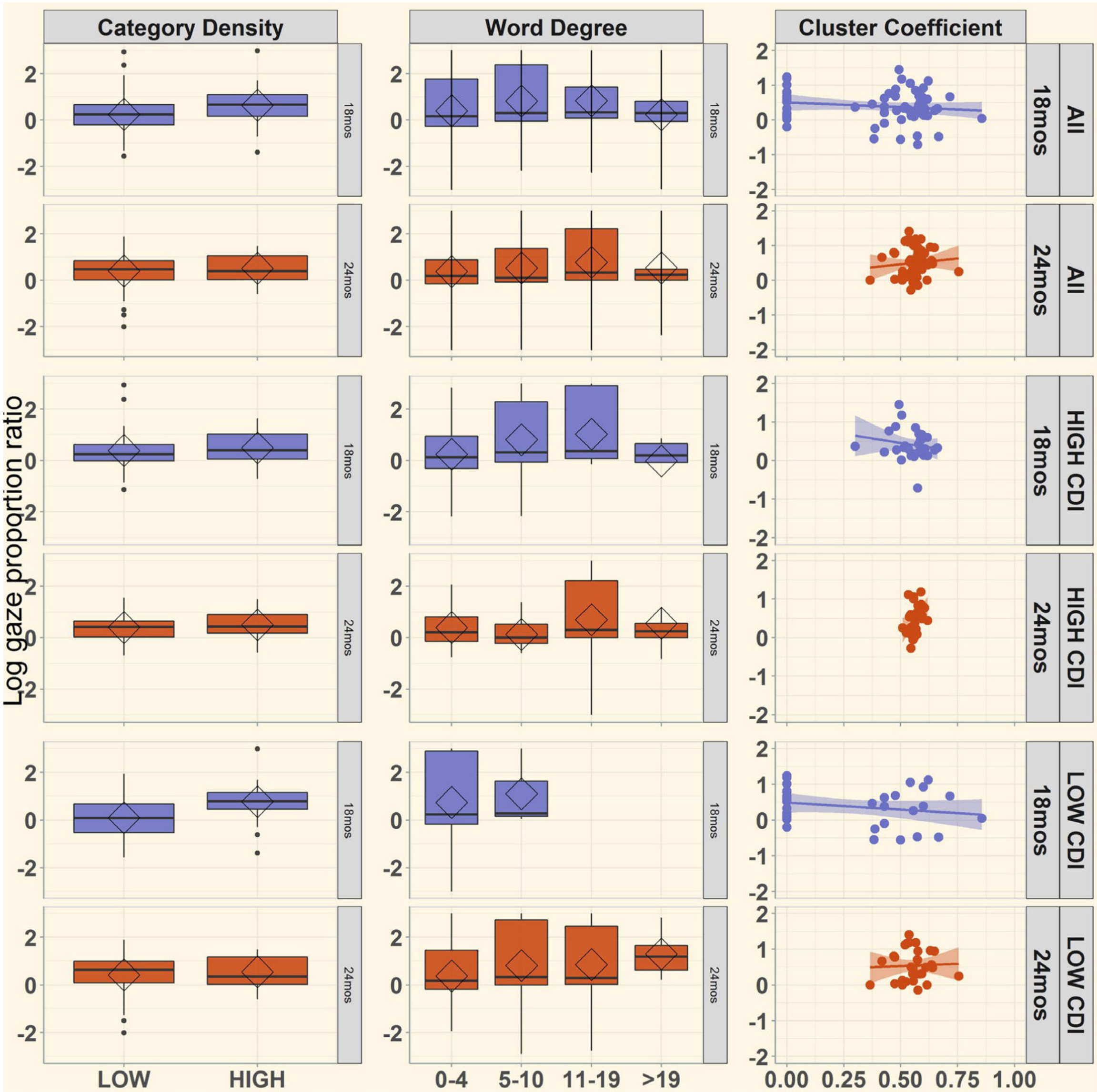


Figure 6.

Relations between semantic structure measures and log-gaze accuracy measures in the semantically unrelated condition at 18 and 24 months. Semantic density is plotted by high and low density categories, word degree is divided according to quartiles of degree distribution, and each point on the cluster coefficient plots represents lexical connectivity for each individual child in the study.

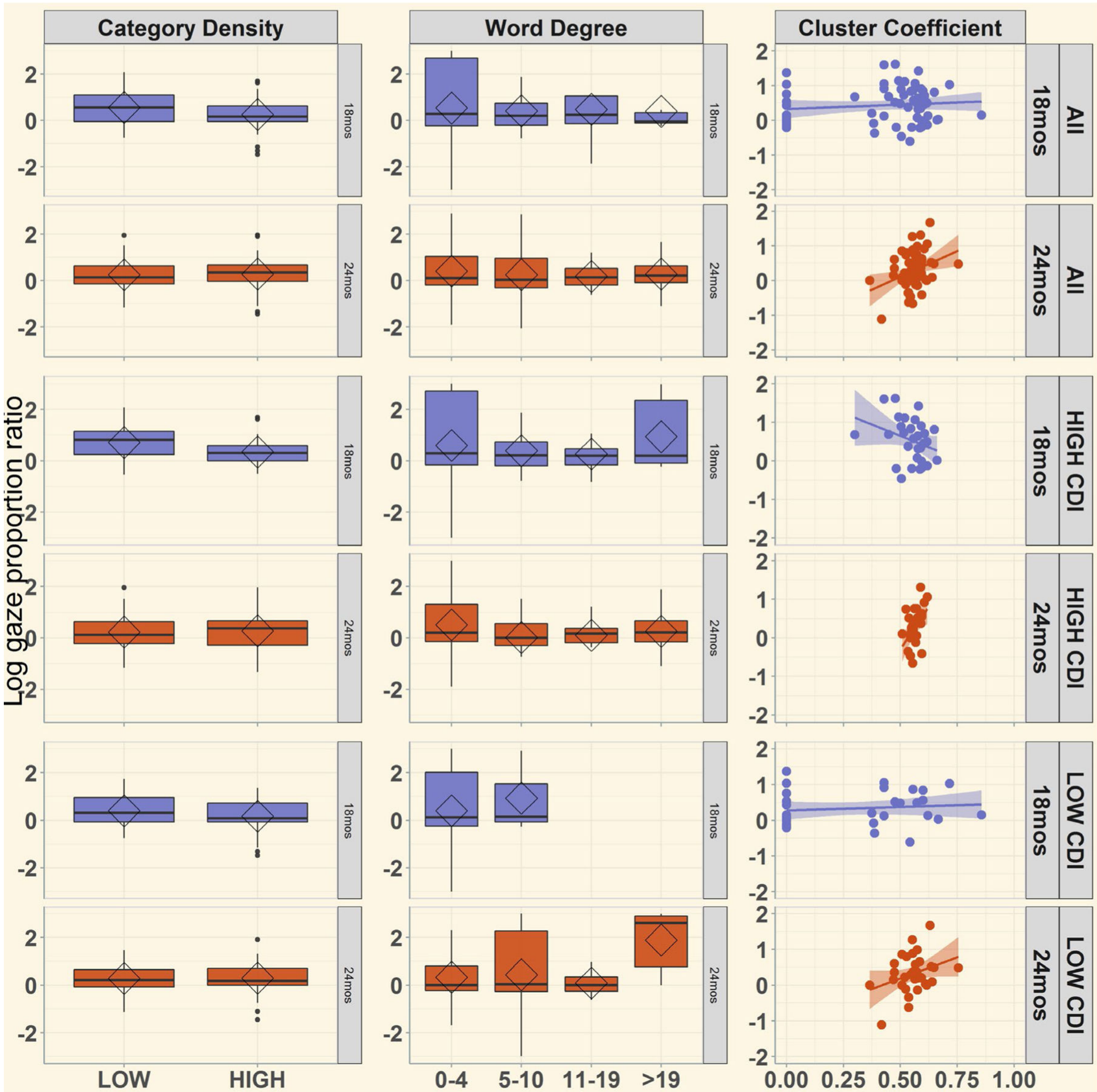


Figure 7.

Relations between semantic structure measures and log-gaze accuracy measures in semantically related trials at 18 and 24 months. Semantic density is box-plotted by high and low density categories, word degree box-plots are sub-divided according to quartiles of degree distribution, and each point on the cluster coefficient plots represents lexical connectivity for each individual child in the study.

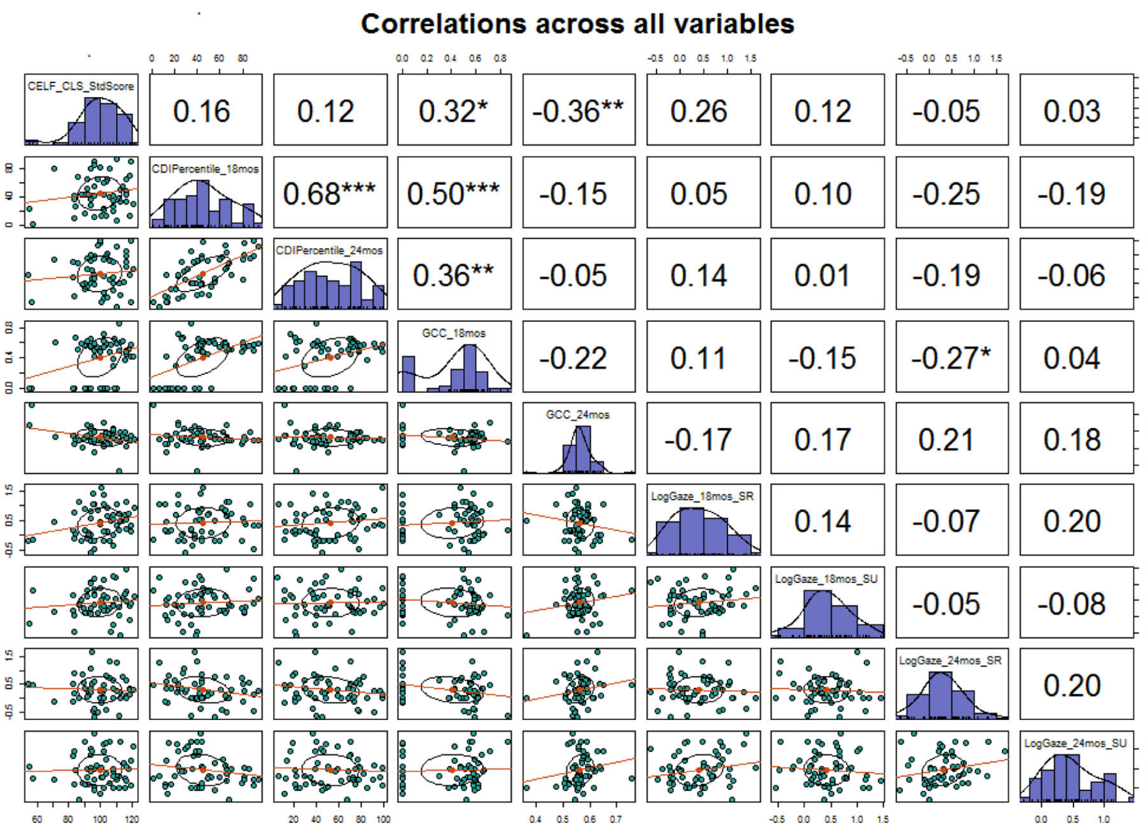


Figure 8.

Pairwise correlation and density distributions of vocabulary size (CDI_Percentile), structure (GCC), and word recognition (LogGaze) in semantically related (SR) and semantically unrelated (SU) trials at 18 and 24 months with CELF standard score (CELF_CLS_StdScore) at 36 months. Lower left trial displays scatterplot with fitted line and ellipses, and values in upper right triangle indicate the correlation coefficient of this relation. Values along the diagonal illustrate the histogram and density estimation of values for each variable. Asterisks for correlation coefficients indicate: * − p<.05, ** − p<.01, *** p<.001

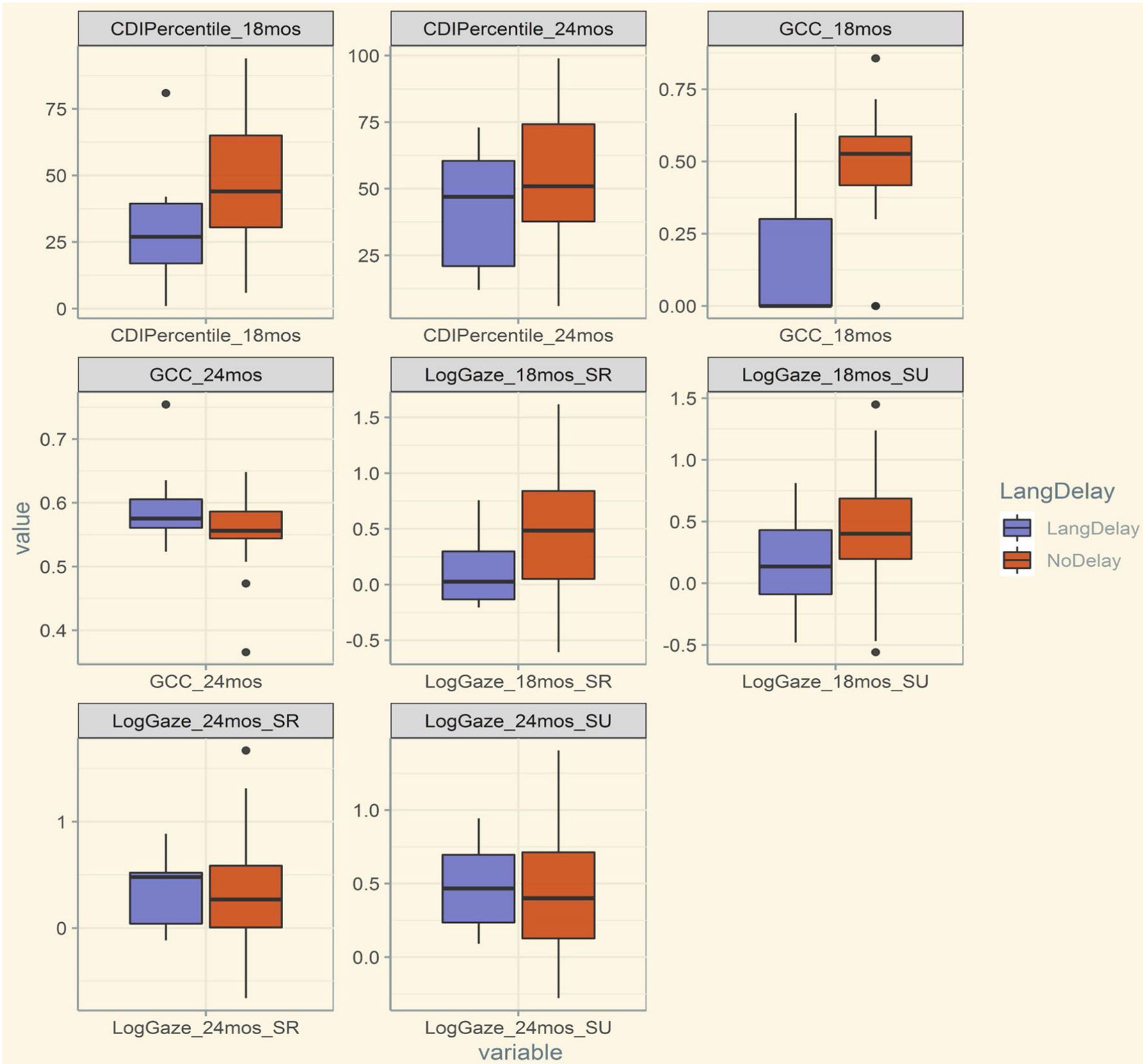


Figure 9.
Boxplots of Vocabulary size, structure, and processing predictor variables as a function of age 3 language delay status, with language delay defined as falling at or below 85 SS on the CELF-P2 CLS, and no delay defined as above 85 SS on this measure.

Table 1.

Participant demographics in Experiment 1 and 2.

	Experiment 1 (N=61)	Experiment 2 (N=49)
Sex		
Female	25 (41%)	19 (38.8%)
Male	36 (59%)	30 (61.2%)
Race		
African American / Black	8 (13.1%)	8 (16.3%)
Caucasian / White	42 (68.9%)	32 (65.3%)
Multiple	3 (4.9%)	3 (6.1%)
Not reported	8 (13.1%)	6 (12.2%)
Ethnicity		
Hispanic	7 (11.4%)	6 (12.2%)
Not Hispanic	53 (86.9%)	42 (85.7%)
Not reported	1 (1.6%)	1 (2.0%)
Maternal Education		
High School	3 (4.9%)	3 (6.1%)
Some College	12 (19.7%)	10 (20.4%)
College	16 (26.2%)	13 (26.5%)
Graduate or Professional	28 (45.9%)	21 (42.9%)
Not reported	2 (3.3%)	2 (4.1%)

Table 2.

Proportion of children in each category with high and low category density at each age tested in Experiment 1.

	18 months		24 months	
	High	**Low**	**High**	**Low**
ANIMALS	.48	.52	.52	.48
BODY PARTS	.61	.39	.68	.32
CLOTHING	.15	.85	.09	.91
DRINK	.52	.48	.58	.42
FRUIT	.66	.34	.69	.31
VEHICLES	.39	.61	.59	.41

Table 3.

LMER results for unrelated trials*.

A) All participants (N=61) **

	Estimate	Std. Error	t-value	p-value
(Intercept)	*0.371*	*0.220*	*1.684*	*0.095*
Vocabulary	0.023	0.147	0.157	0.875
Semantic category density	**0.344**	**0.138**	**2.487**	**0.013**
Word Degree	0.298	0.267	1.118	0.264
Global Cluster Coefficient	−0.099	0.088	−1.122	0.262
Age	0.015	0.244	0.061	0.951
Vocab. x Sem. Category Density	−0.154	0.136	−1.136	0.256
Vocab. x Word Degree	−0.210	0.168	−1.253	0.211
Vocab. x GCC	0.026	0.082	0.320	0.749
Vocab. x Age	−0.014	0.210	−0.068	0.946
Age x Sem. Category Density	−0.245	0.192	−1.277	0.202
Age x Word Degree	−0.314	0.267	−1.175	0.240
Age x GCC	0.366	0.322	1.139	0.255
Cat. Density x Age x Vocab.	0.063	0.194	0.325	0.745
Degree x Age x Vocab.	0.150	0.180	0.833	0.405
GCC x Age x Vocab.	0.146	0.333	0.437	0.662

B) Higher vocabulary group (N=30)

	Estimate	Std. Error	t-value	p-value
(Intercept)	**0.473**	**0.159**	**2.974**	**0.004**
Semantic category density	0.079	0.182	0.432	0.666
Word Degree	*0.235*	*0.137*	*1.718*	*0.087*
Global Cluster Coefficient	−0.214	0.219	−0.979	0.328
Age	−0.431	0.325	−1.325	0.186
Age x Sem. Category Density	−0.038	0.254	−0.150	0.881
Age x Word Degree	−0.201	0.147	−1.370	0.171
Age x GCC	*1.175*	*0.708*	*1.661*	*0.097*

C) Lower vocabulary group (N=31)

	Estimate	Std. Error	t-value	p-value
(Intercept)	0.393	0.542	0.724	0.469
Semantic category density	**0.567**	**0.192**	**2.953**	**0.003**
Word Degree	0.603	0.735	0.820	0.412
Global Cluster Coefficient	*−0.114*	*0.064*	*−1.785*	*0.075*
Age Group	0.092	0.555	0.166	0.868
Age Grp x Sem. Category Density	*−0.508*	*0.278*	*−1.825*	*0.068*
Age Grp x Word Degree	−0.499	0.741	−0.673	0.501
Age Grp x GCC	0.162	0.284	0.569	0.570

Panel A illustrates results for all participants, and panels B and C report model outcomes for higher and lower vocabulary groups, respectively. Significant effects are highlighted in bold, marginal effects in italics. Covariance matrices are appear with the analytic code on osf.io/kt9gs

*Exploratory analyses also found that category pairings did not contribute to the model, and that removing potential late-talkers below the 10th percentile on MBCDI did not influence the results, these analyses can be viewed on osf.io

**Models were also repeated using robust regression, which controls for potential influence of outliers. Statistical patterns are identical and reported on osf.io

Table 4.

LMER results for related trials*.

A) All participants (N=61) **

	Estimate	Std. Error	t-value	p-value
(Intercept)	**0.746**	**0.205**	**3.646**	**0.0003**
Vocabulary	−0.032	0.152	−0.210	0.834
Semantic category density	**−0.316**	**0.143**	**−2.214**	**0.027**
Word Degree	0.277	0.271	1.024	0.306
Global Cluster Coefficient	0.034	0.091	0.370	0.711
Age	**−0.707**	**0.252**	**−2.806**	**0.005**
Vocab. x Sem. Category Density	−0.167	0.139	−1.202	0.230
Vocab. x Word Degree	−0.140	0.174	−0.805	0.421
Vocab. x GCC	−0.003	0.085	−0.031	0.975
Age x Vocab.	−0.183	0.219	−0.836	0.403
Age x Sem. Category Density	**0.398**	**0.199**	**1.996**	**0.046**
Age x Word Degree	−0.324	0.276	−1.172	0.241
Age x GCC	**0.677**	**0.337**	**2.008**	**0.045**
Cat. Density x Age x Vocab.	0.128	0.200	0.638	0.524
Degree x Age x Vocab.	0.153	0.186	0.820	0.412
GCC x Age x Vocab.	0.401	0.346	1.161	0.246

B) Higher vocabulary group (N=30)

	Estimate	Std. Error	t-value	p-value
(Intercept)	**0.823**	**0.154**	**5.343**	**<0.0001**
Semantic category density	**−0.368**	**0.186**	**−1.979**	**0.048**
Word Degree	0.016	0.135	0.117	0.907
Global Cluster Coefficient	*−0.398*	*0.227*	*−1.750*	*0.081*
Age Group	**−1.152**	**0.337**	**−3.416**	**0.0007**
Age Grp x Sem. Category Density	**0.535**	**0.265**	**2.024**	**0.043**
Age Grp x Word Degree	−0.123	0.151	−0.813	0.416
Age Grp x GCC	**1.997**	**0.745**	**2.681**	**0.008**

C) Lower vocabulary group (N=31)

	Estimate	Std. Error	t-value	p-value
(Intercept)	0.907	0.562	1.615	0.107
Semantic category density	−0.148	0.200	−0.742	0.458
Word Degree	0.691	0.776	0.891	0.374
Global Cluster Coefficient	0.031	0.066	0.475	0.635
Age Group	−0.664	0.587	−1.131	0.258
Age Grp x Sem. Category Density	0.171	0.287	0.594	0.553
Age Grp x Word Degree	−0.634	0.787	−0.805	0.421
Age Grp x GCC	0.271	0.291	0.931	0.352

Panel A illustrates results for all participants, and panels B and C include model outcomes for higher and lower vocabulary groups, respectively. Significant effects are highlighted in bold, marginal effects (.05<*p*<.1) in italics. Covariance matrices are posted with the analytic code on osf.io/kt9gs.

* Exploratory analyses also found that category pairings did not contribute to the model, and that removing potential late-talkers below the 10th percentile on MBCDI did not influence the results, these analyses can be viewed on osf.io

** Models were also repeated using robust regression, which controls for potential influence of outliers. Statistical patterns are identical and reported on osf.io

Table 5.

Linear models for 36 month language outcomes.

A) Model 1: Predicting continuous language outcome (CELF-P2 score)				
	Estimate	**Std. Error**	**t-value**	**p-value**
(Intercept)	**140.4**	**21.7**	**6.48**	**<.0001**
Vocabulary Percentile – 18 months	−0.045	0.115	−0.394	0.695
Vocabulary Percentile – 24 months	0.023	0.097	0.243	0.809
GCC – 18 months	**17.358**	**8.523**	**2.037**	**0.048**
GCC – 24 months	**−94.815**	**37.841**	**−2.506**	**0.016**
LogGaze – 18 months – SemRel	3.510	3.665	0.958	0.343
LogGaze – 18 months - SemUnrel	7.371	4.332	1.702	0.096
LogGaze – 24 months - SemRel	3.245	4.031	0.805	0.425
LogGaze – 18 months – SemRel	1.348	4.827	0.279	0.781
B) Model 2: Predicting categorical language delay outcome (CELF-P2 < 85 SS)				
	Estimate	**Std. Error**	**t-value**	**p-value**
(Intercept)	9.789	6.686	1.464	0.143
Vocabulary Percentile – 18 months	−0.041	0.048	−0.855	0.393
Vocabulary Percentile – 24 months	0.023	0.038	0.609	0.543
GCC – 18 months	**5.915**	**2.869**	**2.061**	**0.039**
GCC – 24 months	−18.345	11.504	−1.595	0.111
LogGaze – 18 months – SemRel	0.695	1.539	0.452	0.652
LogGaze – 18 months - SemUnrel	**3.767**	**1.850**	**2.036**	**0.042**
Log Gaze – 24 months - SemRel	1.267	1.346	0.941	0.347
Log Gaze – 24 months - SemUnrel	−0.783	1.729	−0.453	0.651

Significant effects are highlighted in bold.