# Multiple mechanisms of word learning in late talking children: A longitudinal study

Accepted manuscript version for *Journal of Language Speech and Hearing Research.*

**Authors**:

Rachael W. Cheung[1, 2]

Calum Hartley[1]

Padraic Monaghan[1]

[1] Department of Psychology, Lancaster University, Bailrigg, Lancaster, LA1 4YF, United Kingdom

[2] Psychology in Education Research Centre, Department of Education, University of York, York, YO10 5DD, United Kingdom

**Keywords**: late talkers, expressive vocabulary, word learning, nonword repetition, fast mapping, cross-situational word learning, language delay

**Funding**: Leverhulme Trust (RWC, Leverhulme Trust Doctoral Scholar [DS-2014-014]) and the International Centre for Language and Communicative Development (LuCiD) at Lancaster University funded by the Economic and Social Research Council UK (PM and CH) [ES/L008955/1].

**Abstract**

**Purpose:** To identify variability in word learning mechanisms used by late talking children using a longitudinal study design, which may explain variability in late talking children's outcomes.

**Method:** A cohort of typically developing children (*n* = 40) and children who were classified as late talking children at age 2;0 (≤10th percentile on expressive vocabulary, *n* = 21) were followed up at age 3;0 and at age 3;6. We tested the cohort across tasks designed to isolate different mechanisms involved in word learning: encoding and producing spoken forms of words (using a nonword repetition task), identifying referents for words (using a fast mapping task), and learning associations between words and referents (using a cross-situational word learning task).

**Results:** Late talking children had lower accuracy on nonword repetition than typically developing children, despite most of the sample reaching typical ranges for expressive vocabulary at age 3;6. There were no between-group differences in fast mapping and retention accuracy, however, both were predicted by concurrent expressive vocabulary. Late talking children performed less accurately than typically developing children on cross-situational word learning retention trials, despite showing no between-group differences during training trials. Combining performance across all three tasks predicted approximately 45% of the variance in vocabulary outcomes at the last timepoint.

**Conclusions:** Late talking children continue to have deficits in phonological representation that impact their word learning ability and expressive language abilities, but do not show difficulties in fast mapping novel words. Late talking children may also struggle to retain associative information about word-referent mappings. Late talking children thus use some, but not all, word learning mechanisms differently than typically developing children.

## Introduction

Late talking (LT) children fall at or below the 10th percentile for expressive vocabulary compared to typically developing (TD) children at around 2-years-old, despite the absence of concurrent developmental delays or sensory disorders (Desmarais et al., 2008). LT children are at risk of Developmental Language Disorder (DLD; Leonard, 2014; Reilly et al., 2010); however, there are few consistent factors that enable reliable prediction of DLD risk. Expressive vocabulary alone is not a clinically useful predictor of language delay (Law & Roy, 2008; Leonard, 2009), and demographic predictors such as socioeconomic status, family history, and male gender explain only a small amount of variance in outcomes (Dale et al., 2003; Hammer et al., 2017; Hartas, 2011; Henrichs et al., 2011; Lyytinen et al., 2005; Reilly et al., 2010). Furthermore, LT children continue to show persistent variability in language ability throughout toddlerhood and preschool (Dale et al., 2003; Paul, 1993) and may continue score lower than TD children across different language measures (at 8 –17-years-old; Rescorla, 2002, 2005, 2009).

In order to understand how and why outcomes in LT children differ, we must consider whether or not they learn words in a qualitatively different way to TD children. When learning a word for the first time, children must encode and reproduce novel phonological information (as measured by nonword repetition tasks; Coady & Evans, 2008). They must also map the word to its correct referent (referent selection, as measured by in-the-moment processes during fast mapping tasks; Halberda, 2006), and then develop and retain the word-referent association (retention, as measured in cross-situational word learning, CSWL, tasks; McMurray et al., 2012). These processes enable children to both understand (receptive vocabulary) and produce (expressive vocabulary) words later on. However, few studies examine how LT children utilise these mechanisms, and none have investigated how nonword repetition, fast mapping, and CSWL inter-relate with each other, or with vocabulary size. Furthermore, as children build a lexicon, their existing knowledge of words may also impact on how they learn novel ones (Edwards et al., 2004; Stokes, 2010, 2014). Thus, LT children's performance on word learning tasks must also be examined in relation to early and later vocabulary size.

This study seeks address these gaps in the literature by examining longitudinal performance across multiple word learning tasks in a sample of LT children (as compared to TD children).

***Nonword repetition tasks in LT children***

Nonword repetition tasks require children to repeat a list of novel words immediately after a speaker produces them. Such tasks tap into a multitude of language processes that include encoding and reproducing novel phonological information (Coady & Evans, 2008). The literature regarding correlations between nonword repetition and different language measures, including expressive vocabulary, remains somewhat equivocal (see Schwob et al., 2021, for a review). Stokes and Klee (2009) found that children at or below the 16$^{th}$ percentile on the expressive vocabulary at 2 – 2;6-years-old could be identified based on their nonword repetition performance on one-, two-, three-, and four-syllable nonwords. Similarly, Hodges et al. (2017) identified LT children at 2;1 – 2;11-years-old from nonword repetition performance on monosyllabic stimuli of differing consonant complexity. Marini et al. (2017) also reported that LT children aged ~2;6-years-old were impaired on nonword repetition across one- to four-syllable nonwords as compared to TD children. These studies indicate that LT children are characterised by concurrent delays in the immediate storage and reproduction of novel words at time of identification.

However, if LT children continue to show reduced accuracy in nonword repetition, this would suggest that early expressive delay has an enduring impact on children's ability to encode the phonology of novel words. For example, Marini et al. (2017) found that nonword repetition performance at ~2;6-years-old positively correlated with articulation, naming, semantic fluency, and lexical comprehension approximately 11 months later. Studies that test nonword repetition at older ages have found that LT children identified at 2-years-old are also impaired at ~ 2;6- and ~3-years-old (one- to three-syllable nonwords; Rujas et al., 2017), as well as at 4 – 6-years-old (one- to four-syllable nonwords; D'odorico et al., 2007). Broader assessments suggest that LT children identified at 2;1 – 2;11-years-old still struggle to produce speech when tested on their articulation and phonological abilities at 4 – 5-years-old (Neam et al., 2020). However, none of these studies tested concurrent expressive vocabulary at the time of testing. Conversely, others have found no differences on nonword repetition tasks between TD and LT children who have reached typical expressive vocabulary size for their age at 3-years-old (one- to three-syllable nonwords; MacRoy-Higgins & Dalton, 2015) and at 5-years-old (one- to five-syllable nonwords; Petruccelli et al., 2012).

Thus, whilst LT children may initially have impaired nonword repetition, whether they continue to have difficulties remains less certain. Further research is necessary to determine how expressive vocabulary correlates with nonword repetition over time, and also how nonword repetition ability may tie into other mechanisms of word learning.

***Fast mapping in LT children***

Fast mapping tasks assess the ability of children to comprehend or produce novel words immediately after single exposures (Carey & Bartlett, 1978), but unlike nonword repetition, also require accurate referent selection (selecting the correct object that matches the word). Referent selection can be conceptualised as a competitive process between potential word-referent pairs, where fast mapping is driven by cognitive constraints like process-of-elimination or mutual exclusivity (Halberda, 2006; Markman & Wachtel, 1988; McMurray et al., 2012). Mutual exclusivity assumes that each object has only one label – when faced with two objects where one has a known label and one is unfamiliar, children infer that a novel label must refer to the unfamiliar object. TD children are able to respond accurately when comprehending novel words during fast mapping tasks at around 2-years-old (Bion et al., 2013). However, it is not yet clear whether LT children apply the same strategies as TD children during fast mapping.

Based on their nonword repetition task performance, we would expect LT children to be less accurate at *producing* novel words. However, if LT children are able to perform above-chance and equivalent to TD children when tested purely on *comprehension* of fast mapped words, this would suggest that the initial competitive process involved in referent selection is intact, and that referent selection is not necessarily related to early expressive delay. If, however, LT children show reduced accuracy compared to TD children, this would suggest that early expressive delay may be related to receptive fast mapping abilities during referent selection.

Only a few studies have examined fast mapping in LT children. Weismer et al. (2013) found that LT children aged 2;6-years-old scored above-chance at test for comprehension of familiar and novel words. However, in comparison to TD controls, LT children responded less accurately on production of familiar and novel words and on comprehension of novel words, despite being equally able to comprehend familiar words. In a similar task, MacRoy-Higgins et al. (2013) also found that LT children aged ~2-years-old performed less accurately than TD children on comprehension and production of novel words. Rujas et al. (2019) reported that LT children identified at 2;2-years-old struggled to fast map and extend novel words (comprehension) when tested at three timepoints (~2;2-years-old, ~2;9-years-old and ~3;4-years-old), compared to TD children. However, Rujas et al. did not measure concurrent vocabulary throughout their study, meaning it is unclear whether their later timepoints included LT children who had persistent expressive vocabulary delay.

In addition, fast mapping does not necessarily indicate longer term learning and retention of words (Horst & Samuelson, 2008). Thus far, studies of fast mapping in LT children have not tested retention after a delay. Although TD children aged ~2-years-old may retain words that are highlighted using ostensive naming and have multiple repetitions (Munro et al., 2012), TD children show limited ability to retain words using fast mapping following single exposures even at 3 to 4-years-old (Vlach & Sandhofer, 2012). Thus, we would also expect to observe limited retention in both TD and LT children following a single exposure. However, if LT children show less accurate retention in comparison to TD children, this might indicate that the processes underlying retention after fast mapping are related to early expressive delay.

***Cross-situational word learning in LT children***

Acquiring vocabulary stems from not only identifying referents, but also from longer term learning where associations between words and referents are strengthened over time (McMurray et al., 2012; Romberg & Saffran, 2010). In a typical cross-situational word learning (CSWL) task, learners must use statistical information to correctly pair words and referents from trials that contain ambiguous visual objects and auditory labels, by noting when labels and objects co-occur (Yu & Smith, 2007). Infants (Smith & Yu, 2008) and children (Bunce & Scott, 2017; Vlach & DeBrock, 2019) are able to identify correct word-referent pairs during CSWL tasks.

LT children have been found to require more exposures to learn words (MacRoy-Higgins & Dalton, 2015), as have children with DLD (Gray, 2004, 2006; Kan & Windsor, 2010; Rice et al., 1994). Thus, their novel word learning may be increasingly dependent on repeated statistical information than lexical principles that constrain referent selection. If LT children rely on statistical information and repeated exposures to word-referent mappings, this might result in performance on par with TD children during CSWL, but not in fast mapping tasks, where they only have one exposure to a novel word. Studies of CSWL in children with autism spectrum disorder (ASD), a population with significant language difficulties (Eigsti et al., 2011), have not found any differences in CSWL when compared with TD children matched on receptive vocabulary (Hartley et al., 2020; Venker, 2019). These findings indicate that there were no qualitative differences between the populations in how they utilised statistical information.

If CSWL reflects a general cognitive learning mechanism (the ability to extract statistical information and to use process-of-elimination), rather than a language specific mechanism (McMurray et al., 2012; Yu & Smith, 2012), the processes involved in CSWL may also be less dependent on existing vocabulary than those involved in nonword repetition. However, if LT children are impaired in CSWL as a function of limited ability to extract statistical information, this may help characterise why these children appear to have difficulties adding words to their lexicon over time. Ahufinger et al. (2021) found that although both bilingual children with DLD and TD children (8-years-old) performed above chance when selecting the correct referent for a novel word during CSWL, the DLD sample scored significantly less accurately than TD children. However, they tested word-referent mappings immediately after training, rather than referent selection during training, and did not test the retention of words after a delay, meaning it is not possible to distinguish between referent selection and retention ability for these children. Furthermore, no studies to our knowledge have tested task-based CSWL in LT children, nor do CSWL studies typically relate vocabulary to CSWL task performance.

***The present study: integrating word learning mechanisms with development over time***

Overall, these three tasks – nonword repetition, fast mapping, and CSWL – reflect separate key mechanisms that apply to word learning. However, they may apply differently to children according to their individual language abilities, and their role may change over time.

The extant literature regarding LT children leaves a series of open questions concerning their word learning abilities, and where in the process they may have difficulties.

Firstly, although research suggests that LT children are impaired on non-word repetition (Marini et al., 2017; Stokes & Klee 2009), the literature reports mixed results on whether these children continue to show impairments in non-word repetition (D'odorico et al., 2007; Neam et al., 2020). Secondly, studies that examine fast mapping in LT children indicate potential deficits when tested on immediate comprehension and production of novel words (Weismer et al., 2013). However, these are few, often do not test retention after a delay, and do not always report relationships with expressive vocabulary, making it difficult to identify how concurrent expressive vocabulary might predict fast mapping. Thirdly, although LT children benefit from repeated exposures to novel words (MacRoy-Higgins & Dalton, 2015), children with DLD show impairments in statistical learning (Ahufinger et al., 2021), which may also be found in LT children's CSWL. Finally, despite the heterogeneity of LT children, studies do not always account for the trajectory of vocabulary development over time.

We used a longitudinal design to study a cohort of TD children and LT children recruited at 2;0 – 2;5-years-old and followed up at 3;0 – 3;5-years-old and 3;6 – 3;11-years-old. We investigated whether LT children make use of the same strategies as TD children during different processes involved in word learning, examining how both LT status at 2;0 – 2;5-years-old and concurrent expressive vocabulary relate to these stages. Using a repetition task (PSRep Test; Chiat & Roy, 2007), we tested whether LT children show prolonged deficits in their ability to encode and reproduce novel phonological information (nonwords), as well as assessing how intact their phonological representations are for familiar information (real words). Using a fast mapping task that measures comprehension, we tested whether LT children show intact use of mutual exclusivity during referent selection. Using a CSWL task, we tested children's ability to track co-occurrences between words and objects over multiple individually-ambiguous exposures in order to disambiguate correct word-referent associations. We also tested retention following both fast mapping and CSWL, allowing us to identify whether LT children show deficits in the acquisition of novel word-referent pairs after a short delay.

We addressed the following research questions: 1) do LT children exhibit less accurate performance than TD children across different word learning processes, as captured by nonword repetition, fast mapping, and CSWL tasks? 2) Does early and concurrent expressive vocabulary relate to word learning performance in our sample? 3) Can early and later expressive vocabulary size be predicted by variability in word learning mechanisms (exploratory)?

We hypothesised that LT children would demonstrate lower accuracy across all tasks in comparison to TD children at all time points (Ahufinger et al., 2021; Rujas et al., 2017; Weismer et al., 2013). We also hypothesised that higher expressive vocabulary would correlate with more accurate performance across all tasks (e.g. Marini et al., 2017).[1] Finally, by relating past and present vocabulary to tasks that measure different word learning mechanisms, we highlight which processes relating to word learning in LT children may be atypical, and how the trajectory of children's expressive language development may also be affected by these mechanisms.

## Method

### *Participants*

[1] We originally hypothesised that LT children who had not reached the typical expressive vocabulary size range for their age group would perform less accurately, and that LT children who had caught up to their TD peers would perform on par with TD children. However, as all but two LT children reached typical expressive vocabulary size ranges at T2, and we could only test half of the original sample due to COVID-19, resulting in small subgroups, we utilised instead concurrent expressive vocabulary across all participants at T2 and T3.

Ethical approval for the study was obtained from the Lancaster University Faculty of Science and Technology Ethics Committee and from the National Health Service Research Ethics Service. Participants were recruited as part of a longitudinal project that followed-up LT and TD children between the ages of 2;0 – 2;5-years-old to 3;6 – 3;11-years-old. Participants were recruited using flyers from Lancaster University Babylab, via health visitors in the Lancashire local authority, and from nurseries in the local area. Once consent to contact was obtained, parents completed the Oxford-CDI (Hamilton et al., 2000) and were included if they met one of the following criteria for the two groups: TD with an expressive vocabulary score ≥ 25$^{th}$ percentile, or LT with an expressive vocabulary score ≤ 10$^{th}$ percentile. These criteria were chosen to ensure two distinct groups, with the LT criterion consistent with prior literature (Fisher, 2017). Inclusion criteria also included monolingual English, with no medical history of developmental or sensory delays or disorders. Participants were included regardless of receptive vocabulary or the level of parental awareness and concern regarding their children's vocabulary.

A total of 85 families completed the CDI; of these, 24 children were excluded due to the aforementioned criteria. A total of 61 children (40 TD and 21 LT) comprised the final cohort at T1 (Table 1); [2] 59 children were White British, and two children were from UK minority ethnic groups. All families were from mid-socioeconomic status (SES) backgrounds, measured by parental education levels. Of the LT sample, 5 parents were unaware of the child's LT status, 11 were aware and reported some concern, and 5 were aware but unconcerned. The LT sample recruited also had significantly lower receptive vocabulary (*M* = 258 words, *SD* = 93) than the TD sample (*M* = 381 words, *SD* = 41; Table

[2] We did not formally test cognition at T1. However, a history was taken to identify any concerns (i.e. brief birth history; fine/gross motor, sensory, co-ordination, and social milestones; hearing difficulties; previous diagnoses; visits to medical professionals for developmental concerns; referrals to any services). Children were also observed informally throughout warm-up sessions and testing. Data from T4 (Vineland-3 Adaptive Behavior Scales) did not indicate clear developmental concerns at the older age of ~4-years-old (conducted in 75% of original sample).

1), meaning our sample contained children with both receptive and expressive vocabulary deficits. A total of 9 LT children had expressive-only deficits, and 12 had both expressive and receptive deficits (both expressive and receptive vocabulary on the Oxford-CDI ≤ $10^{th}$ percentile, Hamilton et al., 2000; see Supplemental Materials S2 for further information).

Visits occurred at 2;0 – 2;5-years-old -years-old (baseline T1), 12 months from baseline at 3;0 – 3;5-years-old (T2), and 18 months from baseline at 3;6 – 3;11-years-old (T3). At T1 (2;0 – 2;5-years-old), there were 61 participants (40 TD; 21 LT). Between T1 and T2, 3 TD families dropped out of the study permanently (1 family emergency, 2 uncontactable). One TD and one LT family dropped out for T2 (due to pregnancy), but returned in T3.

At T3, 20 TD and 9 LT families were tested before the COVID-19 pandemic, which interrupted data collection. For the rest of cohort at T3, the remaining LT children (8) were prioritised to be tested online on only expressive and receptive vocabulary questionnaires (TD children were not tested online). An additional timepoint (T4; $N = 46$) at 4;0 – 4;6-years-old was later added that could be administered remotely to gain extra information about the cohort. The progression of the study can be seen in Figure 1.

[Insert Figure 1 here]

***Questionnaires***

Caregivers were asked at each visit if any developmental or medical concerns had been noted by themselves or healthcare professionals, including in-between visits, and were specifically asked about hearing difficulties, such as otitis media symptoms.

The *Oxford-CDI* (Hamilton et al., 2000), a parent-report vocabulary questionnaire that measures how many words children say (expressive) and understand (receptive), was used at T1 to confirm participants’ allocation to either the LT or TD group and measure vocabulary.

The *Expressive and Receptive One Word Picture Vocabulary 4th Edition* (EOWPVT-4/ROWPVT-4; (Martin & Brownell, 2011) were used as measures of expressive and receptive vocabulary at T2 and T3, administered by the experimenter. For the EOWPVT-4, children are shown a picture of an object and asked to name it, and for the ROWPVT-4, children are shown four pictures at a time and asked to point to the picture that shows the specific word asked for.

The *Leiter-3* non-verbal Cognitive Battery (Figure Ground, Form Completion, Classification Analogies, Sequential Order; Roid et al., 2013) was used as a measure of non-verbal IQ at T3.

The *Vineland-3 Domain General Parent-Report* questionnaire (Sparrow et al., 2016) was used at T4 as a measure of general functioning.

***Testing session set-up***

The tasks and questionnaires that were administered at each timepoint can be viewed in Table 2. We utilised a mobile testing set-up to maximise retention of participants in the study, with data collection occurring within a room at the lab or at participants' homes. Where testing took place in the home setting, care was taken to ensure a quiet space and clear environment, with just the child and main caregiver present. During testing, the child was seated on one side of a 1 metre fold-out table on a small chair with the caregiver sitting next to them on the floor, and the experimenter was seated on the other side.

Distinct to the current study, participants were also involved in another study investigating symbolic understanding of pictures (Cheung et al., 2022). As this was unrelated to the word learning tasks, the associated task is not further analysed in this study.

***Nonword Repetition Task: The Preschool Repetition (PSRep) Test (Chiat & Roy, 2007)***

***Stimuli***: The PSRep Test contains 18 word and 18 non-words. Words and non-words have an equal number of varying lengths, ranging from 1-syllable to 3-syllables, and are phonologically matched with identical prosodic structure (see Supplemental Materials, S3).

***Procedure:*** The PSRep Test uses live presentation of words and non-words, designed to maximise young children's participation in nonword repetition tasks. Responses are audio-recorded.

***Scoring:*** As per Chiat and Roy (2007), responses were coded for items correct (accuracy) and syllable loss (only accuracy is reported here; for syllable loss scoring and analyses, see Supplemental Materials, S1). Responses were correct if they had all phonemic segments of the target word in the correct order, without any additional phonemes. Consistent or phonetic variant substitutions, such as dentals for alveolars, were acceptable, and allowances were made for regional accents. An independent second rater (blinded to group allocation) coded all responses from recordings, showing good inter-rater reliability (Cohen's $k$ = .89). The total numbers of correct responses for words and non-words were calculated per child as overall accuracy and used as dependent variables for analysis.

***Fast mapping and retention task (Hartley et al., 2019)***

***Stimuli:*** We adapted Hartley et al.'s (2019) fast mapping task for the required age range, reducing the number of trials. Participants had four one-syllable novel words to learn: *lep, darg, terb, yok* (NOUN database; Horst & Hout, 2016). The novel words were randomly paired with four novel objects for each participant. Novel objects were different colours and shapes, but approximately the same size, and familiar objects were common objects checked for familiarity with the parent beforehand (see Supplemental Materials, S3, for stimuli).

***Procedure:*** Participants began with three warm-up trials where they were asked to select a familiar object from an array of three: 'Look! Can you get the [object name]?'. If they responded correctly, they were told: 'Great job! That is the [object name]!' If they responded incorrectly, they were given feedback: 'Actually, this is the [object name]. Can you get the [object name]? Well done, you touched the [object name]!'.

Participants then completed eight referent selection trials (Figure 2a) – four Familiar and four Unfamiliar. For each trial, the experimenter would say: 'Let's look at some new things!' and display a tray with three objects: two familiar and one novel. On Familiar trials, children were asked to select a familiar object ('Can you get the [familiar object]?'. On Unfamiliar trials, children were asked to select the novel object ('Can you get the [novel word]?'). Regardless of the selection made, the experimenter only said: 'Thank you.' The order in which objects were requested was pseudorandomised with the constraint that no more than two trial types of the same kind occurred in a row, and the position of the objects was counterbalanced using a 3x3 Latin Square across participants.

Participants were then given a five-minute break to play with a simple jigsaw puzzle. On return to testing, participants completed 8 retention trials (Figure 2b). For each trial, they were shown three of the novel objects they had learnt words for in the preceding referent selection trials. The experimenter said: 'Look!', and after 3 seconds, they requested one of the novel objects using the corresponding label for that object ('Can you get the [novel name]?'). This repeated until all novel objects had been requested twice. The position of the three objects was pseudorandomised with the constraint that the target did not appear in the same position more than twice in a row. The order in which objects were asked for was counterbalanced across participants using a 3x3 Latin Square.

**Scoring:** Children were scored as correct ('1') or incorrect ('0') per trial by the experimenter during the task based on whether they selected the target object. Total scores for referent selection trials (Familiar and Unfamiliar) and retention trials (Unfamiliar objects only) were calculated per child and used as dependent variables in analysis.

[Insert Figure 2 here]

***Cross-situational word learning task (CSWL; Hartley et al., 2020)***

***Stimuli***: The CSWL task was adapted from Hartley et al. (2020) for the required age range, reducing the number of novel words and trials. The task was coded using PsyScript3 (https://www.lancaster.ac.uk/psychology/research/research-software/), a simple in-house programme that allows experiments to be run in any HTML5-compliant web browser. Stimuli were presented on a Windows 10 SurfacePad Pro touchscreen. There were four two-syllable novel words to learn over 32 trials: *teebu, blicket, fiffin,* and *verdex* (NOUN Database; Horst & Hout, 2016). For each participant, novel words were pseudo-randomly paired with one of four novel objects with different shapes and colours, but similar sizes (see Supplemental Materials, S3).

***Procedure***: Participants began with three warm-up trials, that were not scored, where they saw two familiar objects and were asked: 'Which is the [familiar word]? Touch the [familiar word]'. The warm-up trials repeated until the participant identified the correct referent for each word, before proceeding to the training trials.

On each training trial, participants saw two objects on the screen. A female voice directed them to: 'Look!'. After viewing the pictures for 2.5 seconds, the same voice asked: 'Which is the [novel word]? Touch the [novel word]' (Figure 3a). Each of the four novel word-referent mappings were presented four times; there were 32 trials in total. Each object appeared four times as a target, and four times as a foil. The target appeared an equal number of times on the left and right of the screen, and the order of trials was randomised. When children made their choice by touching the screen, their choice was recorded and the task automatically advanced to the next trial. If they did not make a choice, the experimenter advanced the trial using a hidden asterisk button in the upper right-hand corner of the screen.

The children then had a five-minute break where they played with the examiner using a puzzle, before commencing the retention trials. They began with three warm-up trials where they saw three familiar objects positioned on the left, middle, and right of the screen. The target appeared in each of the three possible locations on one trial, and the trial order was randomised so that children selected targets in each location before the testing trials began.

Children completed eight retention trials – each novel word-object pair was tested on two trials. Three of the objects from the training trials were presented on the screen at a time. After viewing the pictures for 2.5 seconds, the female voice asked: 'Which is the [novel word]? Touch the [novel word]' (Figure 3b). All objects were used four times as foils across the eight trials. The position order was randomised per participant with the constraints that the target object appeared in each position at least twice, and never more than twice in a row. The testing order was also randomised per participant.

**Scoring:** Responses were recorded electronically in real-time using the touchscreen, and coded as correct ('1') or incorrect ('0') based on whether the child's response matched the target. Total items correct for training trials and retention trials per child were used as dependent variables for analysis.

[Insert Figure 3 here]

***Data analysis***

As we were interested in examining how early classification of LT related to performance across different word learning mechanisms, we report here the predictive effect of expressive vocabulary only (see Supplemental Materials, S1, for analyses of receptive vocabulary). Additionally, to allow for capturing the trajectory of vocabulary development over time, we tested the relation between word learning performance and concurrent expressive ability. Across all tasks, data were analysed with T1 expressive vocabulary as a function of population (TD or LT), and T2 and T3 expressive vocabulary as continuous variables, rather than comparing TD children against a LT group that homogenised LT children into those who reached typical expressive vocabulary range and those who did not.

Expressive and receptive vocabulary scores were analysed separately due to high VIF when entered in the same model (>3; Zuur et al., 2010). Data and R code are available (https://osf.io/feg6d/?view_only=26b5bcbe085f4822bbede23a88a87471).

## Results

We first report the study sample characteristics for each timepoint. We then assess the extent to which the LT and TD groups differed on the nonword repetition task, the fast mapping task, and the CSWL task. Finally, we report how the trajectory of expressive vocabulary relates both predictively and concurrently to these word learning mechanisms.

### ***Sample characteristics for T2 and T3***

The final samples for each task can be seen in Tables 3 and 4. At T2 (3;0 – 3;5-years-old), 56 children (36 TD; 20 LT) from the T1 sample participated (Table 3). All but two LT children were above the 10th percentile on the EOWPVT-4, indicating that most of our T2 sample comprised LT children who had reached the typically developing range for expressive vocabulary at time of testing.[3] All 56 completed the fast mapping and retention task at T2. A total of 53 completed the PSRep Test (34 TD; 19 LT); of these, 3 TD children refused to speak nonwords due to shyness and only completed real word stimuli. A total of 2 TD children and 1 LT child were excluded from the PSRep Test due to a high number of incomplete trials (completing less than half of the stimuli). These numbers concerning exclusion and non-responders for nonwords were consistent with Chiat and Roy's (2007) results in the same age group.

At T3 (3;6 – 3;11-years-old), 29 participants (20 TD; 9 LT) were tested before the COVID-19 pandemic ceased face-to-face testing (Table 4). All were tested on the Leiter-3; Welch two-sample *t*-tests showed that the TD children and LT children did not differ significantly in non-verbal IQ at T2 (Table 3) or T3 (Table 4). One TD child and one LT

---

[3] At T3, one child remained at the 10th percentile, and the other child reached above the 10th percentile (for further information, please see Supplemental Materials, S2).

child did not complete the CSWL task due to fussiness. A total of 27 children completed the training trials in the CSWL task (19 TD and 8 LT), all of whom had completed the fast mapping task and the PSRep Test at T2. A further 2 LT children did not complete the CSWL retention trials due to fatigue; 6 LT children and 19 TD children successfully completed the CSWL retention trials. A further 8 LT children completed the EOWPVT-4 and ROWPVT-4 online at T3; of these, all had completed the fast mapping task and 7 had completed the nonword repetition task at T2.

At T4 (4;0 – 4;6-years-old), 46 participants (28 TD; 18 LT) completed the Vineland-3 remotely via video-call or telephone-call during the COVID-19 pandemic. Of the children who took part in T2, LT children (*M* = 94.82, *SD* = 6.43) scored significantly lower than TD children (*M* = 100.56, *SD* = 6.34) on the Vineland-3 Adaptive Behaviour Composite (ABC; Table 3), but not within thresholds indicative of developmental delay (Sparrow et al., 2016). The ABC combines communication, daily living skills, and socialisation subscales; when examined individually (Table 3), the difference between groups appeared to be due to the socialisation subscale, with no difference in communication or daily living skills. There were also no significant group differences on the motor subscale or in maladaptive behaviour. There were no significant group differences in Vineland-3 scores between TD and LT children who took part in T3 (Table 4).

***Differences in word learning mechanisms between LT and TD children***

For analyses between groups and examining relationships between expressive vocabulary and task performance, general linear mixed effects (GLME) models were employed using the function *glmer* from the package *lme4* in R [v1.1.463]. Across all models, we tested fixed effects of population at T1 to determine how LT status related to task accuracy, and effects of concurrent vocabulary (T2 and T3) to determine how this relation might change with vocabulary development. These models were built up progressively, starting with a null model that contained random effects of participant and target word. Fixed effects were added sequentially, with each model compared to the previous best-fitting model using log likelihood comparisons (Barr et al., 2013). Additional fixed effects tested are detailed within each task section.

***Are LT children impaired on nonword repetition?*** We report item correct (accuracy) as our dependent variable; for results and analyses of syllable loss, please see Supplemental Materials (S1).

Overall, LT children were less accurate than TD children on both nonword (LT $M$ = 0.44, $SE$ = 0.03 vs. TD $M$ = 0.76, $SE$ = 0.02) and word items (LT $M$ = 0.51, $SE$ = 0.03 vs. TD $M$ = 0.84, $SE$ = 0.02).

To examine whether children's performance on the nonword repetition task (PSRep Test) differed according to expressive vocabulary, we predicted accuracy (item correct: incorrect = 0, correct = 1) using two GLME analyses, with: 1) population (determined at T1 using CDI; TD = 0, LT = 1), and 2) concurrent expressive vocabulary as fixed effects. These models were tested alongside fixed effects of word length (number of syllables) and word type (word = 0, non-word = 1), with random effects of participant and target word. Random slopes of participant per word did not converge and so were omitted from the model.

The best-fitting model contained fixed effects of population and word length ($\chi^2(2)$ = 12.73, $p$ = .002; Table 5): LT children scored significantly less accurately than TD children ($p$ < .001). All children scored less accurately as word length increased (2-syllables: $p$ = .007; 3-syllables: $p$ <.001). There was no interaction between population and word length, and no effect of word type.

There was also a predictive effect of concurrent expressive vocabulary (T2, EOWPVT-4) on task accuracy. The best-fitting model to the data contained fixed effects of concurrent expressive vocabulary and word length ($\chi^2(2)$ = 12.79, $p$ = .002; Table 5):

accuracy increased with higher expressive vocabulary ($p < .001$), and all children scored less accurately as word length increased (2-syllables: $p = .007$; 3-syllables: $p < .001$). Again, there was no interaction between expressive vocabulary and word length, and no effect of word type.

***Are LT children impaired in fast mapping?*** As a group, LT children performed above chance (33%) on referent selection trials ($M = 0.84$, $SE = 0.04$; $t(19) = 9.55$, $p < .001$), but not on retention trials ($M = 0.33$, $SE = 0.04$; $t(19) = -0.27$, $p = .604$). TD children performed above chance (33%) on referent selection trials ($M = 0.84$, $SE = 0.03$; $t(35) = 15.46$, $p < .001$) and retention trials ($M = 0.40$, $SE = 0.03$; $t(35) = 1.94$, $p = .030$).

To examine whether children's performance on fast mapping differed according to expressive vocabulary, we predicted accuracy (item correct: incorrect = 0, correct = 1) on referent selection and then retention trials using GLME analyses. These models contained: 1) population (determined at T1 using CDI; TD = 0, LT = 1), and 2) concurrent expressive vocabulary as fixed effects. For models that tested retention trial accuracy, we also added a fixed effect of referent selection accuracy, to assess whether accuracy on referent selection trials affected subsequent retention trials for the same item. Random effects of participant and target item were included. As all participants scored at ceiling on familiar trials, we tested only unfamiliar referent selection trials. A model with random slopes of participant per word did not converge (omitted from the final model).

There was no predictive effect of early expressive vocabulary (T1, CDI) on referent selection or retention trials. There was an effect of concurrent expressive vocabulary (T2 EOWPVT-4) on referent selection trials. A model with fixed effects of concurrent expressive vocabulary provided the best fit to the data ($\chi^2(1) = 15.53(1)$, $p$-value $< .001$; Table 6). This showed that participants' accuracy during referent selection trials for unfamiliar words increased with concurrent expressive vocabulary ($p < .001$).

There was also an effect of concurrent vocabulary for retention trials.[4] A model with fixed effects of concurrent expressive vocabulary, referent selection accuracy, and an interaction between expressive vocabulary and referent selection accuracy provided the best fit ($\chi^2(3) = 9.20(3)$, *p*-value = .027; Table 6). This model indicated that higher expressive vocabulary predicted greater accuracy ($p = .023$), and that responding accurately on a referent selection trial significantly increased the likelihood of responding correctly on the corresponding retention trial for the same word ($p = .042$). The interaction also indicated that children with higher concurrent expressive vocabulary were more likely to score accurately even if they were incorrect during referent selection. This suggests that not only did higher concurrent expressive vocabulary predict higher accuracy on referent selection trials and subsequent retention trials, but that it may have also enabled children to map words to referents during retention trials even if they had been wrong previously – i.e. children with higher expressive vocabulary may have been able to 'correct' their previous errors actively during testing. However, the effect of this was not significant ($p = .051$), despite the model providing significantly better fit to the data with the interaction than without, so must be interpreted cautiously.

***Do LT children show impairments in cross-situational word learning?*** Both TD and LT children performed above chance (50%) on training trials (TD: $M = 0.63$, $SE = 0.02$; $t(18) = 3.94$, $p < .001$; LT: $M = 0.62$, $SE = 0.03$; $t(7) = 3.39$, $p = .006$). TD children performed above chance (33%) on retention trials ($M = 0.52$ $SE = 0.04$; $t(18) = 3.87$, $p < .001$), whereas LT did not ($M = 0.31$, $SE = 0.07$; $t(5) = -0.33$, $p = .622$).

[4] A possibility is that the difference in the predictive effect of expressive vocabulary at T1 and T2 was due to a difference in variable type, as T1 was discrete, and T2 was continuous. An additional analysis was run with T1 expressive vocabulary used as a continuous variable, which yielded the same results. Thus, this difference was not likely to be due to variable type.

To examine the effects of early and concurrent expressive vocabulary, we used GLMEs to predict accuracy in training (referent selection) trials and retention trials. We first tested the relation of accuracy to early expressive vocabulary (fixed effect: T1 population, TD or LT, using CDI), then tested relations with concurrent expressive vocabulary (fixed effect: T3, EOWPVT-4), with a random effect of participant. Models with random effects of target item, and random slopes of participant per target item, failed to converge so were omitted.

There were no effects of early or concurrent expressive vocabulary on training trials. However, there was a significant effect of population on retention trial accuracy ($\chi^2(1) = 5.04$, $p = .025$; Table 7): the best fitting model to the data containing a fixed effect of Population (LT or TD) that indicated LT children performed significantly less accurately than TD children.[5] This must be interpreted with caution as only 6 LT children completed this part of the task due to the COVID-19 restrictions limiting data collection. There was no effect of concurrent expressive vocabulary for retention trials.

***Predicting early and later expressive vocabulary from combined mechanisms***

As an exploratory analysis, we used linear models to test the extent to which performance across all tasks combined related to early (T1) and later (T3) vocabulary (*lm* base function in R). This enabled us to determine how the child's developing expressive vocabulary influenced the mechanisms investigated in the tasks.

Using data from all timepoints (T1, T2, T3) from children who completed all three tasks ($N = 22$; 6 LT, 16 TD), the model significantly predicted 32% of variance in children's T1 vocabulary at 2;0 – 2;5-years-old (Table 8; *adjusted* $R^2 = 0.32$; $F(5, 16) = 2.04$; $p = .046$). However, only the PSRep Test was a significant predictor of children's past T1 vocabulary at 2;0 – 2;5-years-old. When predicting future T3 vocabulary at 3;6 – 3;11-years-old, all three

[5] An additional analysis using T1 as a continuous variable (as for the fast mapping task) was not possible, as these models failed to converge.

tasks combined predicted 45% of the variance (Table 8; *adjusted* $R^2 = .45$; $F(5, 16) = 4.45$; $p = .010$). Of the predictor variables, only the PSRep Test predicted future vocabulary.

As the analysis that contained all three timepoints was considerably smaller due to interruption of T3 data collection, we also conducted additional analyses using data from children who completed all tasks at T2 and for whom we had T3 vocabulary data. The model (data: $N = 53$; 19 LT, 34 TD) predicted 40% of the variance in children's past T1 vocabulary at 2;0 – 2;5-years-old (Table 9; *adjusted* $R^2 = 0.40$; $F(3, 49) = 12.32$; $p < .001$). Only the PSRep Test was significant in relation to children's early vocabulary. When relating to T3 vocabulary at 3;0 – 3;5-years-old, the model (data: $N = 33$; 16 LT, 17 TD) predicted 47% of the variance (Table 9; *adjusted* $R^2 = 0.47$; $F(3, 29) = 10.28$; $p < .001$). The PSRep Test and fast mapping retention accuracy predicted children's later vocabulary.

## Discussion

Studies of word learning mechanisms in LT children offer the chance to unpack how early expressive language is delayed and relates to word learning over time. We utilised three critical tasks that highlight key mechanisms involved in word learning: perception and production of phonology, selection and retention of referents, and acquisition of associations between words and referents. We further tested LT and TD children's vocabulary development during the study to investigate how vocabulary growth related to these word learning mechanisms.

### ***LT children continue to show impairments in phonology, but are able to select referents accurately***

LT children were impaired on nonword repetition, consistent with the literature (e.g. Marini et al., 2017). However, unlike Weismer et al. (2013), LT children did not show evidence of impaired performance as compared to TD children during referent selection in fast mapping or in CSWL, with the two groups performing similarly. This may have been as a result of our sample containing older LT children who had reached the typical expressive vocabulary age range, whereas previous studies have tested LT children who remained delayed at a younger age. LT children who have reached typical vocabulary ranges thus appear to be able to fast map unfamiliar words on par with TD children. LT children also scored at ceiling for comprehension of familiar items during fast mapping on par with TD children, but scored less accurately on the PSRep Test for real words as well as for nonwords. This demonstrated that although they were able to identify known objects without difficulty, LT children's ability to produce both familiar and unfamiliar words was compromised.

***LT children show possible deficits in retaining statistical information from the environment***

LT children showed evidence of impairment on CSWL retention trials, but not on fast mapping retention trials. Our results suggest that despite reaching the TD range for expressive vocabulary, LT children may have a weaker encoding of links between words and referents that is tapped by tasks which test retention from repeated exposures, such as CSWL, but not by tasks that test only single exposures and immediate referent selection, such as fast mapping. Although our CSWL results must be interpreted with caution, given the much smaller sample due to COVID-19, they do suggest fertile ground for future research for testing CSWL in language delay.

Of note is that CSWL performance did not relate to concurrent vocabulary across the sample. This lack of relation may be attributed to the reduced sample sizes at T3 due to COVID-19, random variability, or could be task-related. Alternatively, CSWL performance may be secondary to more general cognitive processes that run parallel to vocabulary acquisition, such as working memory, which may also be implicated in nonword repetition (Marini et al. 2017).

Although we did not find differences in fast mapping abilities or in initial CSWL referent selection trials, other studies that directly test online learning have found differences during the learning process itself, despite no differences in overall accuracy. For example, Ellis et al. (2015) tested novel word learning with an eye-tracker (looking-while-listening paradigm) in 18-month-olds and found that, although TD and LT children looked equivalently at the target, there were between-population differences in looking behaviour during testing. They proposed that LTs divided their attention between target and foil equally, being uncertain about the target, whereas TD children predominantly focussed on the target. This is consistent with Ahufinger et al. (2021), who found children with DLD showed more ambivalence when fixating between targets and competitors at test during CSWL, whereas TD children showed a rapid increase in looks to target over competitors. As we did not use eye tracking, it is possible that LTs showed a similar pattern of uncertainty around the target that was not captured by referent selection, but was captured when testing retention trials, which test the robustness of learnt word-referent pairs. It is thus possible that even if accuracy between groups does not differ, the way in which children process information during word learning tasks might.

***Understanding the trajectory of vocabulary development via word learning mechanisms***

Our results also showed how, across the whole sample, children's expressive abilities may interact with word learning mechanisms as their vocabulary develops over time. Firstly, our analyses showed that the higher the concurrent expressive vocabulary of children, the more accurately they scored on not only the PSRep Test, but on both referent selection and retention fast mapping trials. Secondly, both the PSRep Test and fast mapping retention predicted expressive vocabulary scores at the last time point, suggesting that children's ability to not only store phonological information, but also their ability to retain fast mapped word-referent pairs, appears to influence their ability to add words to their expressive vocabulary later on.

Expressive vocabulary may be the result of storing robust semantic *and* phonological representations, where phonological representations are not only auditory and articulatory, but also precise. Although LT children were able to recognise stimuli and activate semantic representations sufficiently during referent selection in order to comprehend novel words during fast mapping and CSWL, they may have had weaker phonological representations stored in their expressive vocabulary secondary to their early language delay. This may have resulted in a reduced ability to produce both words already in the lexicon (real words) and to utilise existing knowledge to produce novel words. Similarly, children with DLD show difficulty encoding novel words, though they are able to consolidate new words on par with TD children once they are encoded, indicating that the precision of phonological representations during the initial encoding of novel words may affect children's ability to build expressive vocabulary (Bishop et al., 2012; Gordon et al., 2021). Notably, although the majority of LT children reached the normal range for TD expressive vocabulary at 3;0- and 3;5-years-old, they still had much lower expressive vocabulary scores than the TD group. Reduced performance in the CSWL retention trials also hint at possible additional deficits in retaining statistical information that may compound LT children's ability to add to their existing lexicon, but require further investigation in a larger sample.

Overall, these results are consistent with Edwards and colleagues (Edwards et al., 2004; Munson et al., 2005) and Stokes (2010, 2014; Stokes et al., 2012) who suggest that, as a part of a dynamic system between phonology and the lexicon, smaller expressive vocabularies result in less support for storing, generating, and using phonological representations, which in turn feeds back into further development of the lexicon. Although both receptive and expressive vocabulary tests tap both phonological and semantic representations, expressive vocabulary places more weight on stored phonological representations that connect both auditory processes (involved in recognising words) and oromotor processes (involved in articulating words; Edwards et al., 2004). For comprehension tasks, phonological representations can be relatively weak – one only needs to recognise a given stimulus to activate semantic representation. For production, however, both phonological and semantic representations must be sufficiently strong to reproduce a stimulus faithfully enough to be recognised by someone else.

Our results also highlight the need to adopt individual differences as part of language acquisition studies, as opposed to grouping children into categories. For example, our sample of LT children had different levels of parental awareness and concern; some research suggests parental concern can be used as an additional predictor of persistent difficulties (e.g. Bishop et al., 2003), but we were not able to do so due to our sample size. However, throughout our analyses, we used mixed effects models that allow for random effects of participant. For LT children in particular, embracing this heterogeneity may explain a large amount of the variance that has yet to be identified. Moves towards this have been made in LT (Fernald & Marchman, 2012; Perry & Kucker, 2019) and TD populations (Samuelson, 2021), but are yet to be widely adopted as standard. Future studies could also employ the use of mixed effects modelling, as well as testing a wide range of linguistic ability and parental concern, to better characterise LT populations and their subsequent outcomes.

***Limitations and future directions***

One major limitation towards the end of the study was the interruption of testing due to the COVID-19 pandemic. This meant that T3 data was incomplete, and non-verbal IQ data could not be collected for the whole sample. This also meant that only eight and six LT children took part in the CSWL training and retention trials respectively. Findings from the CSWL task must thus be interpreted tentatively, and require replication in a much larger sample, potentially with additional eye-tracking to investigate in-task strategies (Ellis et al., 2015). The fact that two LT children and one TD child could not tolerate the retention trials may also have reflected some individual differences in attention that were not controlled for.

We also did not test fast mapping production or generalisation, only comprehension. This was because the third session was particularly long as a result of the Leiter-3, and pilot testing had shown children had trouble tolerating the session even with breaks. Of note is that Weismer et al. (2013) showed that LT children's vocabulary scores and fast mapping performance were inter-domain (expressive vocabulary predicting production, receptive predicting comprehension), whereas TD children were cross-domain (both vocabulary scores predicting both tasks). As expressive vocabulary predicted fast mapping comprehension across our sample, this suggests the LT children tested here were not limited to inter-domain relationships between task and vocabulary.

Another limitation is that our sample consisted of families from mid-high SES backgrounds who had actively signed up for an 18-month longitudinal study on child development. Hence, our findings may not generalise to samples with different demographic features and resources.

Furthermore, as we did not have a comprehensive profile of speech, language, and communication skills for our sample, we are unable to identify whether our participants presented with, or were at risk for, other language difficulties. Although expressive vocabulary size yielded higher correlations with task performance than receptive vocabulary (Supplemental Materials, S1, Table S6), receptive vocabulary is also useful predictor of language outcomes (Fisher, 2017), and our LT sample were also impaired in receptive vocabulary as a group. Given the variation present within the LT population, future longitudinal research will likely benefit from also taking into account the heterogeneity of other language and communication abilities besides expressive vocabulary, such as receptive vocabulary and grammar (Desmarais et al., 2008; Leonard, 2009).

## Conclusion

This study indicates that LT children are impaired across some, but not all, mechanisms involved in word learning. Despite most LT children reaching the typically developing range of expressive vocabulary size at time of testing, they still exhibit significant differences in their ability to encode and repeat words – making more errors when repeating both real words and non-words – even when individual differences were taken into account. Furthermore, nonword repetition ability combined with fast mapping retention ability at age 3 predicted 45% of the variance in later vocabulary outcomes at age 3;6 years across the whole sample. Together, these results are consistent with LT children having weaker phonological representations in models that describe phonological and lexical development as dynamic processes that affect one another (Edwards et al., 2004; Stokes, 2010; 2014). Finally, although LT children did not show any impairment in referent selection, as tested by fast mapping and CSWL tasks, they may be less able to retain information learnt through CSWL – indicating a potentially fruitful area for research with the return of face-to-face testing. Overall, our results add to the evidence base surrounding word learning mechanisms in LT children by highlighting the interplay between expressive vocabulary and word learning mechanisms over time.

**Acknowledgments:** Our gratitude to Peter Tovee for helping us with our mobile set-up and to Professor Jeannette Schaeffer (University of Amsterdam), for useful project discussions. Thank you to Delyth Piper for help with coding data.

## References

Ahufinger, N., Guerra, E., Ferinu, L., Andreu, L., & Sanz-Torrent, M. (2021). Cross-situational statistical learning in children with developmental language disorder.

*Language, Cognition and Neuroscience*, *0*(0), 1–21. https://doi.org/10.1080/23273798.2021.1922723

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bion, R. A. H., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, *126*(1), 39–53. https://doi.org/10.1016/j.cognition.2012.08.008

Bishop, D. V. M., Barry, J. G., & Hardiman, M. J. (2012). Delayed retention of new word-forms is better in children than adults regardless of language ability: A factorial two-way study. *PLoS ONE*, *7*(5), e37326. https://doi.org/10.1371/journal.pone.0037326

Bishop, D. V. M., Price, T. S., Dale, P. S., & Plomin, R. (2003). Outcomes of early language delay: II. Etiology of transient and persistent language difficulties. *Journal of Speech, Language & Hearing Research*, *46*(3), 561–565. https://doi.org/10.1044/1092-4388(2003/045)

Bunce, J. P., & Scott, R. M. (2017). Finding meaning in a noisy world: Exploring the effects of referential ambiguity and competition on 2·5-year-olds' cross-situational word learning. *Journal of Child Language*, *44*(3), 650–676. https://doi.org/10.1017/S0305000916000180

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, *15*, 17–29.

Cheung, R. W., Hartley, C., & Monaghan, P. (2022). Receptive and expressive language ability differentially support symbolic understanding over time: Picture comprehension in late talking and typically developing children. *Journal of*

*Experimental Child Psychology*, *214*, 105305. https://doi.org/10.1016/j.jecp.2021.105305

Chiat, S., & Roy, P. (2007). The Preschool Repetition Test: An evaluation of performance in typically developing and clinically referred children. *Journal of Speech, Language, and Hearing Research*, *50*(2), 429–443. https://doi.org/10.1044/1092-4388(2007/030)

Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders*, *43*(1), 1–40. https://doi.org/10.1080/13682820601116485

Dale, P. S., Price, T. S., Bishop, D. V. M., & Plomin, R. (2003). Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. *Journal of Speech, Language, and Hearing Research: JSLHR*, *46*(3), 544–560.

Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., & Rouleau, N. (2008). Systematic review of the literature on characteristics of late-talking toddlers. *International Journal of Language & Communication Disorders*, *43*(4), 361–389. https://doi.org/10.1080/13682820701546854

D'odorico, L., Assanelli, A., Franco, F., & Jacob, V. (2007). A follow-up study on Italian late talkers: Development of language, short-term memory, phonological awareness, impulsiveness, and attention. *Applied Psycholinguistics*, *28*(1), 157–169. https://doi.org/10.1017/S0142716406070081

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research*, *47*(2), 421–436. https://doi.org/10.1044/1092-4388(2004/034)

Eigsti, I.-M., de Marchena, A. B., Schuh, J. M., & Kelley, E. (2011). Language acquisition in autism spectrum disorders: A developmental review. *Research in Autism Spectrum Disorders*, *5*(2), 681–691. https://doi.org/10.1016/j.rasd.2010.09.001

Ellis, E. M., Borovsky, A., Elman, J. L., & Evans, J. L. (2015). Novel word learning: An eye-tracking study. Are 18-month-old late talkers really different from their typical peers? *Journal of Communication Disorders*, *58*, 143–157. https://doi.org/10.1016/j.jcomdis.2015.06.011

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically-developing and late-talking toddlers. *Child Development*, *83*(1), 203–222. https://doi.org/10.1111/j.1467-8624.2011.01692.x

Fisher, E. L. (2017). A systematic review and meta-analysis of predictors of expressive-language outcomes among late talkers. *Journal of Speech, Language, and Hearing Research*, *60*(10), 2935–2948. https://doi.org/10.1044/2017_JSLHR-L-16-0310

Gordon, K. R., Storkel, H. L., Uglow, S., & Ohlmann, N. B. (2021). Word learning by preschool-age children with Developmental Language Disorder: Impaired encoding and robust consolidation during slow mapping. *Journal of Speech, Language & Hearing Research*, *64*(11), 4250–4270. https://doi.org/10.1044/2021_JSLHR-21-00046

Gray, S. (2004). Word learning by preschoolers with specific Language Impairment: predictors and poor learners. *Journal of Speech, Language & Hearing Research*, *47*(5), 1117–1132. https://doi.org/10.1044/1092-4388(2004/083)

Gray, S. (2006). The relationship between phonological memory, receptive vocabulary, and fast mapping in young children with Specific Language Impairment. *Journal of*

*Speech, Language, and Hearing Research*, *49*(5), 955–969. https://doi.org/10.1044/1092-4388(2006/069)

Halberda, J. (2006). Is this a dax which I see before me? Use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cognitive Psychology*, *53*(4), 310–344. https://doi.org/10.1016/j.cogpsych.2006.04.003

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language*, *27*(3), 689–705. https://doi.org/10.1017/S0305000900004414

Hammer, C. S., Morgan, P., Farkas, G., Hillemeier, M., Bitetti, D., & Maczuga, S. (2017). Late talkers: A population-based study of risk factors and school readiness consequences. *Journal of Speech, Language, and Hearing Research*, *60*(3), 607–626. https://doi.org/10.1044/2016_JSLHR-L-15-0417

Hartas, D. (2011). Families' social backgrounds matter: Socio-economic factors, home learning and young children's language, literacy and social outcomes. *British Educational Research Journal*, *37*(6), 893–914. https://doi.org/10.1080/01411926.2010.506945

Hartley, C., Bird, L.-A., & Monaghan, P. (2019). Investigating the relationship between fast mapping, retention, and generalisation of words in children with autism spectrum disorder and typical development. *Cognition*, *187*, 126–138. https://doi.org/10.1016/j.cognition.2019.03.001

Hartley, C., Bird, L.-A., & Monaghan, P. (2020). Comparing cross-situational word learning, retention, and generalisation in children with autism and typical development. *Cognition*, *200*, 104265. https://doi.org/10.1016/j.cognition.2020.104265

Henrichs, J., Rescorla, L., Schenk, J. J., Schmidt, H. G., Jaddoe, V. W. V., Hofman, A., Raat, H., Verhulst, F. C., & Tiemeier, H. (2011). Examining continuity of early expressive vocabulary development: The generation R study. *Journal of Speech, Language, and Hearing Research: JSLHR*, *54*(3), 854–869. https://doi.org/10.1044/1092-4388(2010/09-0255)

Hodges, R., Munro, N., Baker, E., McGregor, K., & Heard, R. (2017). The Monosyllable Imitation Test for toddlers: Influence of stimulus characteristics on imitation, compliance and diagnostic accuracy. *International Journal of Language & Communication Disorders*, *52*(1), 30–45. https://doi.org/10.1111/1460-6984.12249

Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavior Research Methods*, *48*(4), 1393–1409. https://doi.org/10.3758/s13428-015-0647-3

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, *13*(2), 128–157. https://doi.org/10.1080/15250000701795598

Kan, P. F., & Windsor, J. (2010). Word learning in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, *53*(3), 739–756. https://doi.org/10.1044/1092-4388(2009/08-0248)

Law, J., & Roy, P. (2008). Parental report of infant language skills: A review of the development and application of the Communicative Development Inventories. *Child and Adolescent Mental Health*, *13*(4), 198–206. https://doi.org/10.1111/j.1475-3588.2008.00503.x

Leonard, L. B. (2009). Is expressive language disorder an accurate diagnostic category? *American Journal of Speech-Language Pathology / American Speech-Language-Hearing Association*, *18*(2), 115–123. https://doi.org/10.1044/1058-0360(2008/08-0064)

Leonard, L. B. (2014). Part II: Desrcibing the data: Lingustic and nonlingustic findings: Chapter 5: Exploring the boundaries of SLI. In *Children with Specific Language Impairment* (pp. 151–180). MIT Press.

Lyytinen, P., Eklund, K., & Lyytinen, H. (2005). Language development and literacy skills in late-talking toddlers with and without familial risk for dyslexia. *Annals of Dyslexia*, *55*(2), 166–192. https://doi.org/10.1007/s11881-005-0010-y

MacRoy-Higgins, M., & Dalton, K. P. (2015). The influence of phonotactic probability on nonword repetition and fast mapping in 3-year-olds with a history of expressive language delay. *Journal of Speech, Language, and Hearing Research*, *58*(6), 1773–1779. https://doi.org/10.1044/2015_JSLHR-L-15-0079

MacRoy-Higgins, M., Schwartz, R. G., Shafer, V. L., & Marton, K. (2013). Influence of phonotactic probability/neighbourhood density on lexical learning in late talkers. *International Journal of Language & Communication Disorders / Royal College of Speech & Language Therapists*, *48*(2), 188–199. https://doi.org/10.1111/j.1460-6984.2012.00198.x

Marini, A., Ruffino, M., Sali, M. E., & Molteni, M. (2017). The role of phonological working memory and environmental factors in lexical development in Italian-speaking late talkers: A one-year follow-ip study. *Journal of Speech, Language & Hearing Research*, *60*(12), 3462–3473. https://doi.org/10.1044/2017_JSLHR-L-15-0415

Martin, N., & Brownell, R. (2011). *Expressive and Receptive One-Word Picture Vocabulary Tests–4 (EOWPVT-4/ROWPVT-4)* (4th ed.). Academic Therapy Publications.

McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, *119*(4), 831–877. https://doi.org/10.1037/a0029872

Munro, N., Baker, E., Mcgregor, K., Docking, K., & Arciuli, J. (2012). Why word learning is not fast. *Frontiers in Psychology*, *3*, 41. https://doi.org/10.3389/fpsyg.2012.00041

Munson, B., Edwards, J., & Beckman, M. E. (2005). Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. *Journal of Speech, Language, and Hearing Research*, *48*(1), 61–78. https://doi.org/10.1044/1092-4388(2005/006)

Neam, S. Y., Baker, E., Hodges, R., & Munro, N. (2020). Speech production abilities of 4- to 5-year-old children with and without a history of late talking: The tricky tyrannosaurus. *International Journal of Speech-Language Pathology*, *22*(2), 184–195. https://doi.org/10.1080/17549507.2019.1638968

Paul, R. (1993). Patterns of development in late talkers: Preschool years. *Journal of Childhood Communication Disorders*, *15*(1), 7–14. https://doi.org/10.1177/152574019301500103

Perry, L. K., & Kucker, S. C. (2019). The heterogeneity of word learning biases in late-talking children. *Journal of Speech, Language & Hearing Research*, *62*(3), 554–563. https://doi.org/10.1044/2019_JSLHR-L-ASTM-18-0234

Petruccelli, N., Bavin, E. L., & Bretherton, L. (2012). Children with specific language impairment and resolved late talkers: Working memory profiles at 5 years. *Journal of Speech, Language, and Hearing Research*, *55*(6), 1690–1703. https://doi.org/10.1044/1092-4388(2012/11-0288)

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E., Prior, M., Cini, E., Conway, L., Eadie, P., & Bretherton, L. (2010). Predicting language outcomes at 4 years of age: Findings from Early Language in Victoria Study. *Pediatrics*, peds.2010-0254. https://doi.org/10.1542/peds.2010-0254

Rescorla, L. (2002). Language and reading outcomes to age 9 in late-talking toddlers. *Journal of Speech, Language, and Hearing Research*, *45*(2), 360–371. https://doi.org/10.1044/1092-4388(2002/028)

Rescorla, L. (2005). Age 13 language and reading outcomes in late-talking toddlers. *Journal of Speech, Language, and Hearing Research*, *48*(2), 459–472. https://doi.org/10.1044/1092-4388(2005/031)

Rescorla, L. (2009). Age 17 language and reading outcomes in late-talking toddlers: Support for a dimensional perspective on language delay. *Journal Of Speech, Language, And Hearing Research: JSLHR*, *52*(1), 16–30. https://doi.org/10.1044/1092-4388(2008/07-0171)

Rice, M. L., Oetting, J. B., Marquis, J., Bode, J., & Pae, S. (1994). Frequency of input effects on word comprehension of children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *37*(1), 106–122. https://doi.org/10.1044/jshr.3701.106

Roid, G. H., Miller, L. J., Pomplun, M., & Koch, C. (2013). *Leiter-3: Leiter International Performance Scale* (3rd ed.). Stoelting Company.

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *WIREs Cognitive Science*, *1*(6), 906–914. https://doi.org/10.1002/wcs.78

Rujas, I., Casla, M., Mariscal, S., Lázaro López-Villaseñor, M., & Murillo Sanz, E. (2019). Effects of grammatical category and morphology on fast mapping in typically developing and late talking toddlers. *First Language*, 0142723719828258. https://doi.org/10.1177/0142723719828258

Rujas, I., Mariscal, S., Casla, M., Lázaro, M., & Murillo, E. (2017). Word and nonword repetition abilities in Spanish language: Longitudinal evidence from typically developing and late talking children. *The Spanish Journal of Psychology*, *20*, E72. https://doi.org/10.1017/sjp.2017.69

Samuelson, L. K. (2021). Toward a precision science of word learning: Understanding individual vocabulary pathways. *Child Development Perspectives*, *15*(2), 117–124. https://doi.org/10.1111/cdep.12408

Schwob, S., Eddé, L., Jacquin, L., Leboulanger, M., Picard, M., Oliveira, P. R., & Skoruppa, K. (2021). Using nonword repetition to identify Developmental Language Disorder in monolingual and bilingual children: A systematic review and meta-analysis. *Journal of Speech, Language & Hearing Research*, *64*(9), 3578–3593. https://doi.org/10.1044/2021_JSLHR-20-00552

Smith, L. B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

Stokes, S. F. (2010). Neighborhood density and word frequency predict vocabulary size in toddlers. *Journal of Speech, Language, and Hearing Research*, *53*(3), 670–683. https://doi.org/10.1044/1092-4388(2009/08-0254)

Stokes, S. F. (2014). The impact of phonological neighborhood density on typical and atypical emerging lexicons. *Journal of Child Language*, *41*(3), 634–657. https://doi.org/10.1017/S030500091300010X

Stokes, S. F., Kern, S., & Dos Santos, C. (2012). Extended Statistical Learning as an account for slow vocabulary growth. *Journal of Child Language*, *39*(1), 105–129. https://doi.org/10.1017/S0305000911000031

Stokes, S. F., & Klee, T. (2009). The diagnostic accuracy of a new test of early nonword repetition for differentiating late talking and typically developing children. *Journal Of*

*Speech, Language, And Hearing Research: JSLHR*, *52*(4), 872–882. https://doi.org/10.1044/1092-4388(2009/08-0030)

Venker, C. E. (2019). Cross-situational and ostensive word learning in children with and without autism spectrum disorder. *Cognition*, *183*, 181–191. https://doi.org/10.1016/j.cognition.2018.10.025

Vlach, H. A., & DeBrock, C. A. (2019). Statistics learned are statistics forgotten: Children's retention and retrieval of cross-situational word learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *45*(4), 700–711. https://doi.org/10.1037/xlm0000611

Vlach, H. A., & Sandhofer, C. M. (2012). Fast mapping across time: Memory processes support children's retention of learned words. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00046

Weismer, S. E., Venker, C. E., Evans, J. L., & Moyle, M. J. (2013). Fast mapping in late-talking toddlers. *Applied Psycholinguistics*, *34*(1), 69–89. https://doi.org/10.1017/S0142716411000610

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*(5), 414–420. https://doi.org/10.1111/j.1467-9280.2007.01915.x

Yu, C., & Smith, L. B. (2012). Modeling cross-situational word-referent learning: Prior questions. *Psychological Review*, *119*(1), 21–39. https://doi.org/10.1037/a0026182

Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, *1*(1), 3–14. https://doi.org/10.1111/j.2041-210X.2009.00001.x

**Figure legends:**

**Figure 1. Study diagram showing the progression of longitudinal study and sample sizes across timepoints**.

**Figure 2. Fast mapping and retention task: example of a) referent selection trials; b) retention trials.**

**Figure 3. Cross-situational word learning task: a) example of two training trials: the learner is able to infer that the gasser must be the blue object, based on co-occurrence across the trials; b) example of retention trials.**

**Supplemental Materials:**

**S1 contains syllable loss analyses for the nonword repetition task, descriptive statistics and receptive vocabulary analyses for each task, and correlations of task performance with expressive and receptive vocabulary.**

**S2 contains further information about the LT sample and individual variation in performance for each task.**

**S3 contains further information about each task (approximate time to administer and stimuli).**

**Table 1.**

**Sample demographics at T1.**

| | *Mean (SD)* | | *Welch's two-sample t-test* | | |
|---|---|---|---|---|---|
| | *TD (n = 40)* | *LT (n = 21)* | *t-value (df)* | *p-value* | *Cohen's d* |
| *Age (years;months)* | 2;2 (0;1) | 2;2 (0;1) | - | - | - |
| *Sex (m : f)* | 17: 23 | 14 : 7 | - | - | - |
| *T1 CDI receptive* [a] | 380.98 (40.54) | 257.90 (93.42) | -5.76 (24.02) | <.001 | -1.71 |
| *T1 CDI expressive* [a] | 324.75 (77.89) | 60.05 (49.54) | -16.15 (56.66) | <.001 | -4.06 |

[a] *Raw scores were used*

**Table 2.**

**Measures administered at each timepoint.**

| **Timepoint (years; months)** | **Measures** | **Tasks** |
|---|---|---|
| *T1: 2;0 – 2;5-years-old (N = 61)* | Oxford-CDI | |
| *T2: 3;0 – 3;5-years-old (N = 56)* | EOWPVT-4<br>ROWPVT-4 | Fast mapping and retention<br>PSRep Test |
| *T3: 3;6 – 3;11-years-old* [a] *(N = 29)* | EOWPVT-4<br>ROWPVT-4<br>Leiter-3 | Cross-situational word learning |
| *T4: 4;0 – 4;6-years-old (N = 46; remote testing)* | Vineland-3 | |

[a] *An additional 8 children were also tested remotely at this time point, on only the EOWPVT-4 and the ROWPVT-4 during COVID-19.*

**Table 3. Sample demographics and vocabulary at T2.**

| | *Mean (SD)* | | *Welch's two-sample t-test* | | |
|---|---|---|---|---|---|
| | *TD (n = 36)* | *LT (n = 20)* | *t-value (df)* | *p-value* | *Cohen's d* |
| *Age (years;months)* | 3;2 (0;2) | 3;2 (0;2) | - | - | - |
| *Sex (m : f)* | 16 : 20 | 14 : 6 | - | - | - |
| *T1 CDI receptive* [a] | 381.17 (42.11) | 265.80 (88.37) | -5.50 (23.89) | <.001 | -1.67 |
| *T1 CDI expressive* [a] | 324.97 (79.56) | 61.60 (50.30) | -15.15 (52.97) | <.001 | -3.96 |
| *T2 ROWPVT-4* | 116.03 (10.48) | 108.60 (10.16) | -2.59 (40.41) | .013 | -0.72 |
| *T2 EOWPVT-4* | 120.47 (9.21) | 107.75 (13.80) | -3.69 (28.61) | <.001 | -1.08 |
| | *TD (n = 19)* | *LT (n = 9)* | | | |
| *T3 Non-verbal IQ (Leiter-3)* | 98.63 (6.86) | 92.00 (12.81) | -1.46 (10.24) | .175 | -0.65 |
| | *TD (n = 25)* | *LT (n = 17)* | | | |
| *T4 Vineland ABC* [b] | 100.56 (6.34) | 94.82 (6.43) | -2.85 (34.16) | .007 | -0.90 |
| *T4 Vineland Com* | 144.16 (180.54) | 103.47 (25.47) | -1.11 (25.39) | .278 | -0.32 |
| *T4 Vineland DLS* | 96.28 (5.43) | 94.76 (7.87) | -0.69 (26.20) | .496 | -0.22 |
| *T4 Vineland Soc* | 98.48 (7.56) | 92.82 (7.94) | -2.31 (33.36) | .027 | -0.73 |
| *T4 Vineland Mot* | 98.28 (6.01) | 95.76 (8.99) | -1.01 (25.63) | .322 | -0.33 |
| *T4 Vineland MB* [a] | 6.00 (3.04) | 7.12 (4.39) | 0.91 (26.30) | .370 | 0.30 |

*NB: unless otherwise specified, standardised scores were used.*

*ABC = Adaptive Behaviour Composite; Com = Communication subscale; DLS = Daily Living Score subscale; MB = Maladaptive Behaviour subscale; Mot = Motor subscale; LT = late talking; PSRep = Preschool Repetition; Soc = Socialisation subscale; TD = typically developing; vocab = vocabulary*

[a] *Raw scores used*

[b] *This is a composite of Communication, Daily Living Skills, and Socialisation subscales*

**Table 4. Sample demographics and vocabulary at T3.**

| | *Mean (SD)* | | *Welch's two sample t-test* | | |
|---|---|---|---|---|---|
| | *TD (n = 19)* | *LT (n = 8)* | *t-value (df)* | *p-value* | *Cohen's d* |
| *Age (years;months)* | 3;9 (0;1) | 3;9 (0;2) | - | - | - |
| *Sex (m : f)* | 7 :12 | 6: 2 | - | - | - |
| *T1 CDI receptive* [a] | 394.32 (23.41) | 275.38 (58.60) | -5.56 (7.96) | <.001 | -2.67 |
| *T1 CDI expressive* [a] | 350.11 (67.95) | 74.38 (57.95) | -10.71 (15.44) | <.001 | -4.37 |
| *T3 ROWPVT-4* | 119.68 (5.39) | 111.88 (7.68) | -2.62 (10.03) | .026 | -1.18 |
| *T3 EOWPVT-4* | 122.53 (8.60) | 108.38 (13.73) | -2.70 (9.41) | .023 | -1.24 |
| *T3 Non-verbal IQ (Leiter-3)* | 98.84 (6.74) | 93.75 (12.49) | -1.09 (8.77) | .305 | -0.51 |
| | *TD (n = 15)* | *LT (n = 8)* | | | |
| *T4 Vineland ABC* [b] | 102.07 (6.32) | 97.63 (6.42) | -1.59 (14.22) | .134 | -0.70 |
| *T4 Vineland Com* | 110.2 (6.20) | 113.0 (34.61) | 0.23 (7.24) | .827 | 0.11 |
| *T4 Vineland DLS* | 97.80 (4.54) | 97.25 (6.36) | -0.22 (10.91) | .832 | -0.10 |
| *T4 Vineland Soc* | 98.27 (8.97) | 96.25 (7.74) | -0.56 (16.41) | .581 | -0.24 |
| *T4 Vineland Mot* | 99.13 (6.31) | 100.13 (9.37) | 0.27 (10.49) | .794 | 0.12 |
| *T4 Vineland MB* [a] | 6.00 (3.34) | 6.25 (5.28) | 0.12 (10.07) | .906 | 0.06 |

*NB: unless otherwise specified, standardised scores were used.*

*ABC = Adaptive Behaviour Composite; Com = Communication; DLS = Daily Living Score; MB = Maladaptive Behaviour; Mot = Motor subscale; LT = late talking; PSRep = Preschool Repetition; Soc = Socialisation subscale; TD = typically developing*

[a] *Raw scores used*

[b] *This is a composite of Communication, Daily Living Skills, and Socialisation subscales*

**Table 5.**

**Nonword repetition task: general linear mixed effects model summary results of best-fitting model predicting item correct by fixed effects of T1 and T2 expressive vocabulary.**

| Relation with early expressive vocabulary (measured at T1: 2;0 – 2;5-years-old) | | | | |
|---|---|---|---|---|
| *Fixed effect* | *estimate* | *SE* | *z-value* | *p-value* |
| *(intercept)*[a] | 2.90 | 0.39 | 7.39 | < .001 |
| T1 population (late talking) | -2.14 | 0.36 | -6.03 | < .001 |
| 2-syllable words | -1.22 | 0.46 | -2.68 | .007 |
| 3-syllable words | -1.73 | 0.46 | -3.78 | < .001 |
| Relation with concurrent expressive vocabulary (measured at T2: 3;0 – 3;5-years-old) | | | | |
| *Fixed effect* | *estimate* | *SE* | *z-value* | *p-value* |
| *(intercept)* | -7.55 | 1.79 | -4.21 | < .001 |
| T2 expressive (EOWPVT-4)[a] | 8.32 | 1.52 | 5.47 | < .001 |
| 2-syllable words | -1.22 | 0.46 | -2.68 | .007 |
| 3-syllable words | -1.73 | 0.46 | -3.79 | < .001 |

[a] *Rescaled using x/100 to allow model fit*

**Table 6.**

**Fast mapping task: summary results of best-fitting general linear mixed effects model predicting accuracy in referent selection and retention trials by concurrent expressive vocabulary.**

| Referent selection trial accuracy | | | | |
|---|---|---|---|---|
| *Fixed effect* | *estimate* | *SE* | *z-value* | *p-value* |
| *(Intercept)* | -3.07 | 2.12 | -1.45 | .148 |
| T2 expressive vocabulary (EOWPVT)[a] | 6.36 | 1.66 | 3.83 | <.001 |

| Retention trial accuracy | | | | |
|---|---|---|---|---|
| *Fixed effect* | *estimate* | *SE* | *z-value* | *p-value* |
| *(intercept)* | -8.14 | 3.23 | -2.52 | .012 |
| T2 expressive vocabulary (EOWPVT)[a] | 6.45 | 2.84 | 2.28 | .023 |
| Referent selection (correct) | 6.96 | 3.43 | 2.03 | .042 |
| T2 expressive[a] * referent selection (correct) | -5.86 | 3.00 | -1.96 | .051 |

[a] *Rescaled using x/100 to allow model fit*

**Table 7.**

**Cross-situational word learning task: summary results of best-fitting general linear mixed effects model predicting accuracy in retention trials with early expressive vocabulary (T1, CDI).**

| *Fixed effect* | *estimate* | *SE* | *z-value* | *p-value* |
|---|---|---|---|---|
| *(intercept)* | 0.08 | 0.18 | 0.45 | .653 |
| T1 population (LT) | -0.89 | 0.39 | -2.28 | .022 |

**Table 8.**

**Predicting early and later expressive vocabulary by task performance (accuracy) using data from all timepoints (T1, T2, T3; *N* = 22).**

| Predicting early (T1; CDI) expressive vocabulary at 2;0 – 2;5-years-old | | | | |
|---|---|---|---|---|
| *Variance* | *estimate* | *SE* | *t-value* | *p-value* |
| *(intercept)* | -44.56 | 211.70 | -0.21 | 0.836 |
| Preschool Repetition Test | 4.41 | 1.49 | 2.96 | .009 |
| Fast mapping referent selection | -0.75 | 1.60 | -0.47 | .645 |
| Fast mapping retention | 0.90 | 1.47 | 0.61 | .549 |
| Cross-situational word learning referent selection | -0.14 | 3.15 | -0.05 | .964 |
| Cross-situational word learning retention | 0.86 | 1.48 | 0.58 | .570 |
| Predicting later (T3; EOWPVT-4) expressive vocabulary at 3;6 – 3;11-years-old | | | | |
| *Variance* | *estimate* | *SE* | *t-value* | *p-value* |
| *(intercept)* | 53.07 | 19.00 | 2.79 | .013 |
| Preschool Repetition Test | 0.32 | 0.13 | 2.43 | .027 |
| Fast mapping referent selection | 0.01 | 0.14 | 0.04 | .968 |
| Fast mapping retention | 0.23 | 0.13 | 1.77 | .096 |
| Cross-situational word learning referent selection | 0.52 | 0.28 | 1.85 | .083 |
| Cross-situational word learning retention | -0.03 | 0.13 | -0.25 | .080 |

**Table 9.**

**Predicting early and later expressive vocabulary by task performance (accuracy) using data from completed timepoints (T1, T2).**

| Predicting early expressive vocabulary at 2;0 – 2;5-years-old (T1, CDI; N = 53) | | | | |
|---|---|---|---|---|
| *Variance* | *estimate* | *SE* | *t-value* | *p-value* |
| (intercept) | 26.75 | 76.30 | 0.35 | 0.727 |
| Preschool Repetition Test | 4.17 | 0.70 | 5.93 | <.001 |
| Fast mapping referent selection | -1.07 | 0.78 | -1.37 | .178 |
| Fast mapping retention | 0.14 | 0.80 | 0.18 | .860 |
| Predicting later expressive vocabulary at 3;0 – 3;5-years-old (T2, EOWPVT-4, N = 33) | | | | |
| *Variance* | *estimate* | *SE* | *t-value* | *p-value* |
| *(intercept*) | 85.92 | 7.92 | 10.84 | <.001 |
| Preschool Repetition Test | 0.28 | 0.07 | 3.99 | <.001 |
| Fast mapping referent selection | 0.04 | 0.09 | 0.48 | .634 |
| Fast mapping retention | 0.22 | 0.09 | 2.53 | .017 |