Supplemental Material S1. Additional analyses (syllable loss analyses for the nonword repetition task, receptive vocabulary analyses for each task, and correlations of task performance with expressive and receptive vocabulary).

A number of analyses were undertaken in addition to those reported in the manuscript. These were as follows:

1) Additional analyses of the nonword repetition task (the PS Rep Test):
    a. The effect of receptive vocabulary on accuracy
    b. The effect of expressive vocabulary on syllable loss
    c. The effect of receptive vocabulary on syllable loss
2) The effect of receptive vocabulary on the fast mapping task
3) The effect of receptive vocabulary on the cross-situational word learning (CSWL) task
4) Cross-correlations of all three tasks (nonword repetition, fast mapping, and CSWL) with children's expressive and receptive vocabulary over time.

The results of these analyses are reported here. For all linear mixed effects (LME) and general linear mixed effects models (GLME) reported, the same procedure was utilised to build models as detailed in the main manuscript. Stimuli used for the nonword repetition task, fast mapping task, and CSWL task can be found in Supplemental Material S3.

1) Nonword repetition task: the Preschool Repetition (PSRep) Test (Chiat & Roy, 2007)

The PSRep Test has two measures of performance: items correct (accuracy) and syllable loss. Main effects of expressive vocabulary on accuracy were reported in the manuscript, however, analyses of syllable loss and effects of receptive vocabulary are reported here. Syllable loss was coded where either two syllables were coalesced, combining the consonant from one and the vowel from the other, or when a vowel was omitted with or without adjacent consonants (apart from /blun/ and /plis/ which are considered acceptable). An independent second rater coded all response for syllable loss, showing good inter-rater reliability (Cohen's k = .86).

Table S1 shows descriptive statistics for the nonword repetition task (PS Rep Test) with a breakdown of item correct (accuracy) and syllable loss across groups, word type, and word length.

Table S1. Nonword repetition task: accuracy and syllable loss mean and standard error.

Word length	**Non words: mean (*SE*)**			
	Late talking (n = 19)		***Typically developing (n = 31)***	
	Accuracy	*Syllable loss*	*Accuracy*	*Syllable loss*
One-syllable words	0.61 (0.05)	*n.a.*	0.91 (0.02)	*n.a.*
Two-syllable words	0.41 (0.05)	0.10 (0.03)	0.70 (0.03)	0.06 (0.02)

Three-syllable words	0.31 (0.04)	0.32 (0.05)	0.67 (0.03)	0.09 (0.02)
	Real words: mean (*SE*)			
	Late talking (n = 19)		***Typically developing (n = 31)***	
	Accuracy	*Syllable loss*	*Accuracy*	*Syllable loss*
One-syllable words	0.63 (0.05)	*n.a.*	0.92 (0.02)	*n.a.*
Two-syllable words	0.49 (0.05)	0.18 (0.05)	0.85 (0.03)	0.04 (0.01)
Three-syllable words	0.41 (0.05)	0.40 (0.05)	0.76 (0.03)	0.14 (0.03)

Note. n.a. = not applicable.

a) Receptive vocabulary and nonword repetition task accuracy: We predicted accuracy (item correct) using two GLME analyses, with fixed effects of 1) T1 receptive vocabulary (CDI), and 2) T2 receptive vocabulary (ROWPVT-4 score). Each model also had fixed effects of word length (number of syllables) and word type (word = 0, nonword = 1), and random effects of participant and target item. Random slopes of participant per target were attempted but caused non-convergence, so were omitted from the model.

There was a significant effect of T1 receptive vocabulary on accuracy. The best-fitting model contained an effect of T1 receptive vocabulary ($\chi^2(2) = 12.71$, $p = .002$; Table S2), indicating that the higher children's receptive vocabularies, the more accurately they scored ($p < .001$), and a fixed effect of word length, indicating that children scored less accurately with when words were longer (2-syllable, $p = .008$; 3-syllable, $p < .001$). There was no effect of word type.

There was also an effect of T2 receptive vocabulary on accuracy. The best-fitting model to the data contained fixed effects of T2 receptive vocabulary, word length, word type, and an interaction between receptive vocabulary and word type, and an interaction between receptive vocabulary and word length ($\chi^2(2) = 6.50$, $p = .039$; Table S2). This model indicated that children with higher receptive vocabularies scored more accurately ($p = .006$). The interaction term between receptive vocabulary and word type indicated that children with higher receptive vocabularies scored less accurately on nonwords than real words ($p = .034$), and the interaction between receptive vocabulary and word length indicated that children with higher receptive vocabularies scored less accurately on 2-syllable words as compared to 1-syllable words, although this was not significant ($p = .050$).

Table S2. Nonword repetition task: linear mixed effects results of best-fitting predicting accuracy by fixed effects of T1 and T2 receptive vocabulary.

Relation with T1 receptive vocabulary at 2;0 – 2;5-years-old				
Fixed effect	*estimate*	*SE*	*z-value*	*p-value*
(intercept)	–1.43	0.92	–1.55	.120
T1 receptive vocabulary (CDI)[a]	1.04	0.25	4.18	< .001
2-syllable words	–1.22	0.46	–2.66	.008

2-syllable words	–1.73	0.46	–3.78	< .001
Relation with T2 receptive vocabulary at 3;0 – 3;5-years-old				
Fixed effect	*estimate*	*SE*	*z-value*	*p-value*
(intercept)	–7.71	2.60	–2.97	.003
T2 receptive vocabulary (ROWPVT-4)[a]	8.97	2.29	3.92	< .001
2-syllable words	2.33	1.85	1.26	.209
3-syllable words	–2.14	1.95	–1.10	.272
Word type (nonword, 1)	2.52	1.49	1.69	.091
T2 receptive × 2-syllables	–3.17	1.61	–1.96	.050
T2 receptive × 3-syllables	0.33	1.69	0.19	.847
T2 receptive × word type (nonword, 1)	–2.73	1.29	–2.12	.034

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

b) Expressive vocabulary and nonword repetition task syllable loss: For the PSRep Test, expressive vocabulary was also tested with regard to syllable loss (only accuracy was reported in the main manuscript).

We predicted syllable loss using two LME analyses, with fixed effects of 1) T1 Population (0 = TD, 1 = LT) and 2) T2 expressive vocabulary (EOWPVT-4 score). Each model also had a fixed effect of word type (real word = 0, nonword = 1) and random effects of participant and target item.

The best-fitting model to the data contained a fixed effect of T1 Population on syllable loss ($\chi^2(1) = 16.56$, $p < .001$; Table S3), indicating that LT children lost more syllables than TD children ($p < .001$), with no effect of word type.

There was also an effect of T2 expressive vocabulary on syllable loss. The best-fitting model to the data contained fixed effects of T2 expressive vocabulary, trial type, and an interaction between expressive vocabulary and trial type, with random effects of participant and target ($\chi^2(2) = 9.43$, $p = .009$; Table S3). This model indicated that those with higher vocabularies lost fewer syllables ($p < .001$) and that all children lost fewer syllables on non-word items in comparison to word items ($p = .002$). The interaction term indicated that children who had higher expressive vocabularies lost more syllables in non-word trials in comparison to word trials ($p = .003$).

Thus, children identified as LT at T1 lost more syllables, despite all but one having reached typical vocabulary size for their age using expressive percentile criteria at time of testing (T2). Those with higher concurrent (T2) expressive vocabularies also lost fewer syllables, particularly in word trials.

Table S3. Nonword repetition task: linear mixed effects results of best-fitting model predicting syllable loss by fixed effects of T1 and T2 expressive vocabulary.

Relation with T1 expressive vocabulary at 2;0 – 2;5-years-old					
Fixed effect	*estimate*	*SE*	*t-value*	*df*	*p-value*
(intercept – typically developing)	0.05	0.02	2.28	65.39	.026

T1 population (CDI; late talking, 1)	0.11	0.03	4.41	52.13	< .001
Relation with T2 expressive vocabulary at 3;0 – 3;5-years-old					
Fixed effect	*estimate*	*SE*	*t-value*	*df*	*p-value*
(intercept)	0.86	0.13	6.46	90.05	< .001
T2 expressive vocabulary (EOWPVT-4)[a]	–0.64	0.11	–5.79	83.53	< .001
Trial type (non-word, 1)	–0.39	0.13	–3.06	1371.00	.002
T2 expressive[a] * trial type (non-word, 1)	0.32	0.10	3.01	1698.79	.003

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

c) Receptive vocabulary and nonword repetition task syllable loss: We predicted syllable loss using GLME analyses with fixed effects of T1 receptive vocabulary (CDI), and T2 receptive vocabulary (ROWPVT-4 score), and word type (word or non-word), and random effects of participant and target item. Random slopes of participant per target were attempted but caused non-convergence, so were omitted from the model.

There was a significant predictive effect of T1 receptive vocabulary on nonword repetition syllable loss. The best-fitting model to the data contained fixed effects of receptive vocabulary, with random effects of participant and target word ($\chi^2(1) = 5.05$, $p = .025$; Table S4). Children with higher receptive vocabularies had lower rates of syllable loss ($p = .025$). There was no interaction between receptive vocabulary and word length.

There was also a predictive effect of T2 receptive vocabulary on syllable loss by way of an interaction with word length. The best-fitting model to the data contained fixed effects of T2 receptive vocabulary and word length, with an interaction between receptive vocabulary and word length, and random effects of participant and target word ($\chi^2(2) = 46.88$, $p < .001$; Table S4). This model indicated that children lost significantly more syllables with increased word length, and the interaction indicated that children with lower receptive vocabularies lost significantly more syllables on three-syllable words ($p < .001$).

Table S4. Nonword repetition task: general linear mixed effects results of best-fitting model predicting syllable loss by fixed effects of T1 and T2 receptive vocabulary.

Relation of accuracy with T1 receptive vocabulary at 2;0 – 2;5-years-old					
Fixed effect	*estimate*	*SE*	*t-value*	*df*	*p-value*
(intercept)	0.23	0.06	3.67	60.53	< .001
T1 receptive vocabulary (CDI)[a]	–0.04	0.02	–2.31	51.39	.025
Relation of accuracy with T2 receptive vocabulary at 3;0 – 3.5-years-old					
Fixed effect	*estimate*	*SE*	*t-value*	*df*	*p-value*
(intercept)	–0.0004	0.17	–0.003	112.1	.998
T2 receptive vocabulary (ROWPVT-4) [a]	–0.0003	0.15	–0.002	109.4	.999
2-syllable words	0.34	0.16	2.08	169.0	.038
3-syllable words	1.27	0.16	7.77	169.0	< .001

T2 receptive vocab × 2-syllable words	–0.23	0.14	–1.60	168.7	.110
T2 receptive vocab × 3-syllable words	–0.94	0.14	–6.62	168.7	< .001

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

2) Receptive vocabulary and the fast mapping task (Hartley et al., 2019)

Referent selection: We predicted referent selection accuracy using two GLME analyses with fixed effects of 1) T1 receptive vocabulary (CDI score), and 1) T2 receptive vocabulary (ROWPVT-4). These models also had random effects of participant and target item. Random slopes of participant per target did not converge, and so were omitted. Only unfamiliar items were used in the models, as all children were at ceiling for familiar items.

There was no effect of T1 receptive vocabulary on referent selection accuracy. There was, however, an effect of T2 receptive vocabulary. A model with fixed effects of T2 receptive vocabulary provided the best fit to the data ($\chi^2(1) = 17.67(1)$, p-value < .001; Table S5). This showed that the higher participants' receptive vocabulary, the more accurately they scored on unfamiliar referent selection trials ($p < .001$).

Retention: We predicted retention accuracy using two GLME analyses with fixed effects of 1) T1 receptive vocabulary (CDI score), and 1) T2 receptive vocabulary (ROWPVT-4). These models also had a fixed effect of previous referent selection accuracy for the same word (incorrect = 0, correct = 1), and random effects of participant and target item. Random slopes of participant per target did not converge, and so were omitted. There was no effect of T1 or T2 receptive vocabulary, yielding no significant improvements in fit over the null model.

In sum, although receptive vocabulary at T1 did not predict fast mapping proficiency, having concurrent higher receptive vocabulary at T2 predicted accuracy on referent selection trials. When data from referent selection and retention trials were combined, higher receptive vocabulary also predicted performance across the task and within retention trials. The effect of concurrent receptive vocabulary on retention trial accuracy was smaller than concurrent expressive vocabulary.

Table S5. Fast mapping task: general linear mixed effects results of best-fitting model predicting accuracy in referent selection trials by fixed effects of T2 receptive vocabulary.

Relation with T2 receptive vocabulary at 3;0 – 3;5-years-old				
Fixed effect	*estimate*	*SE*	*z-value*	*p-value*
(intercept)	–6.81	1.94	–3.51	< .001
T2 receptive vocabulary (ROWPVT-4)[a]	7.67	1.78	4.31	< .001

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

3) Receptive vocabulary and CSWL task (Hartley et al., 2020)

We predicted training trial accuracy and retention accuracy using two GLME analyses with fixed effects of 1) T1 receptive vocabulary (CDI score), and 1) T2 receptive vocabulary (ROWPVT-4). These models also had random effects of participant and target item. Random slopes of participant per target did not converge, and so were omitted. There was no effect of T1 or T2 receptive vocabulary on accuracy.

4) Correlations between task performance (nonword repetition, fast mapping, and CSWL) and receptive and expressive vocabulary over time

We conducted Kendall's rank correlation *tau-b* values (one-tailed) to assess the relationships between word learning task performance and vocabulary over time (Table S6).

At T1 (2;0 – 2;5-years-old), both receptive and expressive vocabulary significantly correlated with PSRep Test accuracy, with expressive vocabulary (τ = 0.49) yielding a higher correlation than receptive (τ = 0.36).

At T2 (3;0 – 3;5-years-old), expressive vocabulary had higher significant correlations than receptive vocabulary on the PSRep Test (expressive: τ = 0.35; receptive: τ = 0.29) and fast mapping referent selection (expressive*:* τ = 0.40; receptive: τ = 0.35).

At T3 (3;6 – 3;11-years-old), expressive vocabulary yielded higher significant correlations than receptive vocabulary on the PSRep Test (expressive: τ = 0.45; receptive: τ = 0.31) and fast mapping retention (expressive: τ = 0.31; receptive: τ = 0.27). Expressive vocabulary also correlated significantly with fast mapping referent selection performance (τ = 0.31), whereas receptive vocabulary did not.

Table S6. Kendall's rank tau (τ)correlations between tasks and vocabulary over time**.**

Vocabulary	**Nonword repetition (PSRep Test)**	**Fast mapping**		**CSWL**	
	Accuracy	*Referent selection*	*Retention*	*Referent selection*	*Retention*
		T1: 2;0 – 2;5-years-old			
Expressive (Oxford-CDI)	**τ = 0.49, p < .001**	τ = 0.06, p = *n.s.*	τ = 0.11, p = *n.s*	τ = –0.10, p = *n.s*	τ = 0.08, p = *n.s*
Receptive (Oxford-CDI)	**τ = 0.36, p < .001**	τ = 0.134, p = *n.s*	τ = 0.05, p = *n.s*	τ = –0.02, p = *n.s*	τ = 0.05, p = *n.s*
		T2: 3;0 – 3;5-years-old			
Expressive (EOWPVT-4)	**τ = 0.35, p < .001**	**τ = 0.40, p < .001**	τ = 0.20, p = .050	τ = 0.08, p = *n.s*	τ = 0.09, p = *n.s*

Receptive (ROWPVT-4)	**τ = 0.29, p = .003**	**τ = 0.35, p = .001**	τ = 0.10, *p = n.s*	τ = 0.19, *p = n.s*	τ = 0.18, *p = n.s*
T3: 3;6 – 3;11-years-old					
Expressive (EOWPVT-4)	**τ = 0.45, p < .001**	**τ = 0.31, p = .023**	**τ = 0.31, p = .014**	τ = 0.13, *p = n.s*	τ = 0.29, *p = n.s*
Receptive (ROWPVT-4)	**τ = 0.31, p = .010**	τ = 0.20, *p = n.s*	**τ = 0.27, p = .035**	τ = 0.12, *p = n.s*	τ = 0.20, *p = n.s*

CDI = Communicative Development Inventories; CSWL = Cross-situational word learning; E/ROWPVT-4 = Expressive/Receptive One Word Picture Vocabulary Test 4th Edition; PSRep = Preschool Repetition; T = timepoint