Two-year-old children’s processing of two-word sequences occurring 19+ times per million and their influence on subsequent word learning

Keywords: word learning, vocabulary acquisition, statistical learning, multi-word sequences, language development.

Abstract

While we know 8-month-olds track the statistical properties of a series of syllables and 2- to 3-year-olds process familiar phrases more efficiently than unfamiliar phrases, less is known about the intermediary level of two-word sequences. In study 1, two-year-old children (N = 45, mean age 651 days) heard two-word sequences consisting of a prime word followed by a noun, with two pictures appearing on screen (depicting the noun and a distractor). Eye-tracking showed that children looked more quickly at the noun picture for two-word sequences occurring an average of 19 times per million and 206 times per million in child-directed speech than for novel sequences. In Study 2, corpus analyses showed that two-year-old children's noun learning increased in line with the frequency of the two-word utterance that preceded it in caregiver speech. This effect holds even after controlling nouns for frequency in caregiver speech, phonemic length, neighborhood density, phonotactic probability and concreteness and after removing nouns produced in isolation by caregivers and nouns produced by children before being produced by caregivers. These studies show that young children's language processing is facilitated by known two-word sequences, allowing the child to focus on more novel aspects of the utterance. Such efficiencies are far-reaching because nearly two-thirds of child-directed utterances contain two-word sequences with frequencies of 19+ per million.

Introduction

Language learning is one of the most important skills for the young child to acquire, with proficiency in language at an early age predicting success in reading and academic achievement (e.g., Bleses, Makransky, Dale, Højen Ari, 2016). We know that 8-month-old infants are able to track the statistical properties of the syllabic components of a linguistic stimulus, using this information to identify probable word boundaries within a speech stream (Saffran, Aslin & Newport, 1996). At two years of age, children are sensitive to the statistical properties of the lexical components of their linguistic input, being more likely to accurately complete familiar four-word sequences than unfamiliar ones (Bannard & Matthews, 2008). A significant determinant of language learning is therefore an ability to capitalize on the statistical properties that are inherent within the language stimulus itself.

Language operates at various levels, from the individual sounds that combine to form words to syntactic structures that combine words to form phrases and utterances. The linguistic information that is present at each level show evidence of statistical learning. For example, children more accurately produce sounds and sound sequences that occur frequently in their language input compared to those that rarely occur (Zamuner, 2009); word learning is more likely to involve lexical items involving those frequent sound sequences (Jones, Cabiddu, Andrews & Rowland, 2020); and children are more able to repeat combinations of known words that form frequently occurring phrases than when they form infrequent phrases (Bannard & Matthews, 2008).

The most parsimonious explanation for statistical effects that operate on different levels of linguistic granularity is the same statistical learning mechanism operating on long-term representations that gradually change in grain size over time (for a review, see Arnon & Snider, 2010; for example implementations of such a mechanism, see Jones & Rowland, 2017; Jones & Macken, 2018). The consequence of an increasing knowledge-base of familiar linguistic information is that children are more able to efficiently process that information. For example, children with larger vocabularies process familiar phrases more quickly (Hurtado, Marchman & Fernald, 2008); and the time taken to look towards one of two noun pictures occurs more quickly with greater familiarity of the phrase involving the noun (Fernald, Pinto, Swingley, Weinberg & McRoberts, 1998). In summary, there are clear language processing efficiencies relating to familiarity gained from frequent exposure to a linguistic stimulus, a theoretical viewpoint that can be traced at least back to Case's (1978) view of automaticity and more generally to usage-based accounts of language (e.g., Bybee, 1998; Ellis, 2002).

Processing efficiency has been examined in some depth at the lexical and phrasal levels. There are extremely robust effects of frequency on the time required by adults to recognise a word (e.g., Andrews, 1989; Howes & Solomon, 1951), and similar effects seem present in young children. For example, 18- to 21-month-olds are just as quick to recognize familiar words when hearing only the initial part of the word compared to hearing the word in its entirety (Fernald, Swingley & Pinto, 2001). Children are also more proficient in their spoken word processing as they get older (Ojima, Matsuba-Kurita, Nakamura & Hagiwara, 2011), a finding that could plausibly be accounted for by age-related increases in word familiarity. More recently, research has shown that familiarity also plays a clear role in processing phrases, with words being recognized more quickly by 18-month-old children when appearing at the end of a familiar sentence frame (e.g., *where's the baby*) than when appearing in isolation (e.g., Fernald & Hurtado, 2006). For 40-month-old children, familiar phrases (e.g., *a lot of noise*) are repeated more quickly than less familiar phrases (e.g., *a lot of juice*) even when the phrases are identical apart from the final, frequency-matched word (Bannard & Matthews, 2008). Again, a plausible explanation for these effects is increased familiarity - in this case to the particular phrases - given that children with greater language exposure respond more quickly to familiar phrases (e.g., Weisleder & Fernald, 2013).

The current paper focuses on three unanswered questions in relation to processing efficiency and familiarity. First, given the effects of familiarity on both words and phrases, it follows that similar effects may be present for word bigrams. We define a word bigram as a sequence of two consecutive words in a spoken utterance. There is evidence that children hold knowledge relating to information that is between the word-level and phrase-level. For example, 3-year-old children use both the word and its gender marking to help identify picture referents of familiar Spanish words (e.g., *la pelota* [feminine article + feminine noun, *ball*]) (Lew-Williams & Fernald, 2007); and children produce fewer errors when generating questions from a high frequency frame (e.g., wh-word + auxiliary + variable) than when a frame is not present (Rowland, 2007).

Second, the vast majority of studies involving young children have used familiar words and phrases that frequently occur in child-directed speech; yet familiarity is a question of degree (see also Jones & Macken, 2015). Given that 9- to 15-month-old infants can hear half a million words in a 3-week period (Swingley, 2007), words and phrases that are far less common may also be considered familiar. In adult word recognition, the facilitatory effect of frequency greatly diminishes as frequency increases - for example, the magnitude of the frequency effect between words of 1 and 2 occurrences per million are the same as those between words of 100 and 200 occurrences per million and between words of 1000 and 2000 occurrences per million (Brysbaert, Mandera & Keuleers, 2018). Thus far, work involving familiar phrases has used frequencies averaging at least 131 per million[1]; relative to other phrases, this likely constitutes very high frequency given that 80% of word types have a frequency per million of fewer than 1 (van Heuven, Mandera, Keuleers & Brysbaert, 2014). We therefore need to establish whether processing efficiencies are seen for word bigrams that are substantially less frequent than the phrases used in previous work involving young children.

Third, we need to determine whether processing efficiencies involving word bigrams - should they exist - influence subsequent word learning. Phrase-level research suggests that processing efficiencies gained from familiar phrases offer greater opportunity to learn new vocabulary items. For example, 18-month-old children who are faster to recognize words appearing at the end of a familiar sentence frame (e.g., *where's the doggy?*) have larger vocabularies one year later (Fernald & Marchman, 2012). Similarly, those 3-year-old children who are able to quickly process a familiar linguistic phrase also have larger vocabularies one year later (Mahr & Edwards, 2018). More directly, 30-month-old children who efficiently process familiar phrases are more likely to learn novel words that appear in familiar sentence frames (Lany, 2018).

The purpose of the current paper is to examine two-year-old children's processing of word bigrams that vary in frequency within child-directed speech to see whether (a) there are processing efficiencies related to the word bigrams, and (b) whether these efficiencies influence subsequent vocabulary learning. In Study 1, we use eye-tracking to assess whether children process the second word in a bigram more quickly when the bigram has been attested in child-directed speech (using two sets of word bigrams with frequencies of 19 times

---

[1] Fernald et al. (1998) bigrams and frequencies in our corpus (for details, see Method, Study 1): the baby (363), the ball (86), the shoe (13), the doggy/doggie (60). Bannard and Matthews (2008): lowest log frequency of 3.62 = raw frequency ~3,390 with ~4 word bigrams per utterance giving ~1.28 million word bigrams; frequency = 3,390 / 1.28 = 2,568 per million.

per million and 206 times per million) than when it has not (zero frequency, or control condition). Based on the word recognition literature, we may expect that the time to process the second word of a bigram decreases as the frequency of the bigram increases. In Study 2, we use corpus analyses to examine whether children's vocabulary acquisition is influenced by the frequency of the word bigrams in utterances that contain novel vocabulary items. Here, we may expect word learning to increase in line with the frequency of the constituent word bigrams. We therefore hope to establish whether processing efficiencies exist in young children even for relatively infrequent word bigrams and if so, how the processing of many utterances (and the subsequent word learning from those utterances) is going to be influenced by word bigram knowledge.

Study 1 – Young children's eye tracking of word bigrams varying in frequency within maternal utterances

Method

Participants

Forty-five 17- to 27-month-old children participated in the study (age range 511-802 days, mean = 651 days, SD = 87, 23 female). All carers reported their children had typical births and normal vision and 7% reported their child as having current or past hearing problems (these were included in the analyses). 78% of carers were educated to college level or above and all reported that English was the main language spoken and heard at home.

Design

There were two experimental conditions (word bigrams that average 19 occurrences per million and 206 occurrences per million in child-directed speech, hereafter labelled 19-frequency and 206-frequency) together with a control condition containing novel, or zero frequency bigrams having the same second word (target) as that of the 19-frequency and 206-frequency conditions. We also included age as a factor, since exposure to word bigrams would be expected to increase with age. The dependent measure was the time taken to fixate on a target picture that corresponded to the second word of each bigram.

Processing of two-word sequences

Materials

Word bigram frequencies were derived from maternal utterances using the same 16 child-directed transcriptions used by Jones and Rowland (2017)[2]. These consisted of 12 children from the Manchester corpus (Theakston, Lieven, Pine, & Rowland, 2001) containing 34-hour transcriptions from audiotaped interactions between mothers and children, taken twice every three weeks over a period of one year. At the beginning of the study, child age was 22 months (range 20-24 months). Four additional children of the same age as the children in the Manchester corpus were added: Eleanor and Fraser from the MPA-EVA corpus, Thomas from the Thomas corpus (for both, see Lieven, Salomo, & Tomasello, 2009) and Lara from the Lara corpus (Rowland & Fletcher, 2006; all corpora available on CHILDES, see MacWhinney, 2000). Table 1 shows the characteristics of each dataset.

Table 1. Characteristics of the child-directed speech transcripts (child names are pseudonyms).

Child	Utterances	Word tokens	Word bigrams
Anne	33,388	124,534	91,146
Aran	33,638	167,684	134,046
Becky	24,696	88,172	63,476
Carl	21,540	78,853	57,313
Dominic	32,862	119,086	86,224
Eleanor	15,446	52,866	37,420
Fraser	22,329	79,038	56,709
Gail	24,873	94,137	69,264
Joel	24,935	94,035	69,100
John	17,594	72,438	54,844
Lara	21,694	75,137	53,443
Liz	17,742	69,809	52,067
Nic	27,304	105,029	77,725
Ruth	33,668	127,424	93,756
Thomas	25,818	130,068	104,250
Warren	23,781	117,536	93,755

---

[2] Note we used the same random samples as per Jones and Rowland (2017) for the additional four children so that sample size across all children was approximately equal.

Mean (SD)	*25,081* *(5,877)*	*99,740* *(29,600)*	*74,658* *(24,623)*
Total	*401,308*	*1,595,846*	*1,194,538*

Stimuli were word bigrams where the first word was the prime and the second the target. All targets (and distractors, see later) were imageable nouns that appeared at least 400 times across all transcripts and appeared in every transcript at least once. Eight experimental word bigram prime-target pairs (e.g., *poor dolly*) were selected having a mean frequency of 206 per million (the 206-frequency condition) and appearing in an average of 14 of the 16 maternal transcripts. Eight experimental word bigram prime-target pairs (e.g., *pretty dress*) were selected having a mean frequency of 19 per million (the 19-frequency condition) and appearing in an average of 8 of the 16 maternal transcripts. Sixteen control word bigrams were created by pairing plausible primes with the previously determined targets (e.g., *start dolly, safe dress*) and confirming that the word bigrams were not present in any of the transcripts. Distractor images were shown alongside target images ensuring that the prime-distractor word bigram was plausible in spoken English but did not appear in any of the transcripts (e.g., *poor train, pretty cow*). Table 2 illustrates the stimuli together with relevant descriptive statistics[3]. We should note that the constraints on finding suitable word bigrams meant the targets in the 19-frequency and 206-frequency conditions were close to being significantly different in their word frequency (19-frequency M = 775.50, SD = 342.11, 206-frequency M = 1,272.88, SD = 582.11, $t(18) = 2.08$, $p = .056$). That said, the 206-frequency targets are on average phonetically longer and have more syllables, though again, neither are significant (phonemes: 19-frequency M = 2.88, SD = .41, 206-frequency M = 3.50, SD = .57, $t(18) = 1.78$, $p = .096$; syllables: 19-frequency M = 1.00, SD = 0.00, 206-frequency M = 1.25, SD = .21, $t(18) = 1.53$, $p = .149$). We also provide the frequencies of each bigram and target word in each maternal transcript in Table 3. We note that there is some variability for a small number of bigrams (primarily *big boy* and *good girl*) that presumably relate to the gender of the child.

All prime words, target words, and control words were recorded individually together with a set of attention words (*oh!, hey!, aww!, ahh!*) by a female native English speaker. Note

---

[3] Note that due to the difficulty of finding a complete set of word bigrams that fulfilled all our criteria, some constraints were minimally relaxed for a small number of word bigrams, as shown in Table 2.

that recording the words individually was necessary in order for the experimental and control conditions to be the same; one could expect any effect seen to be stronger in continuous speech (e.g., because of co-articulation across the two words). The images used were colored, easily recognizable prototypical pictures of the labels used as target and distractor.

Table 2. Prime-target and prime-distractor details for 206-frequency and 19-frequency word bigrams together with prime-target controls (frequencies are total frequency across all transcripts). Matched components of stimulus items are in bold. For distractors, the prime is the same as per the target; for controls it is the target words that are identical. Number of mothers indicates how many maternal transcripts contain the word bigram.

Prime-Target (P-T)	Prime-Distractor (P-D)	Control	Word bigram frequency		Number of mothers		Word frequency		Number of valid RT trials	
			P-T	P-D	P-T	P-D	Target	Distractor	P-T	Control
				High familiarity condition						
big boy	**big** sheep	just **boy**	67	0	10	0	1,520	408	10	2
in bed	**in** teddy	even **bed**	105	1	16	1	982	636	13	2
poor dolly	**poor** train	start **dolly**	67	2	11	2	823	1954	8	5
some milk	**some** car	still **milk**	114	0	15	0	598	2373	9	5
that box	**that** head	called **box**	120	2	16	2	1,528	1026	2	12
your hand	**your** monkey	better **hand**	329	1	16	1	700	492	5	5
good girl	**good** chair	keep **girl**	1,074	0	13	0	1,828	662	9	9
little baby	**little** cake	never **baby**	89	0	13	0	2,204	673	7	3
		Mean	*246 (206 per mill.)*	*1*	*14*	*1*	*1,273*	*1,028*	*8* *(N = 63)*	*5* *(N = 43)*
				Low familiarity condition						
new book	**new** fish	mixing **book**	19	0	12	0	1,204	605	14	15
next door	**next** bridge	mean **door**	35	0	8	0	821	849	8	1
pretty dress	**pretty** cow	safe **dress**	21	0	7	0	547	679	9	4
four eggs	**four** feet	making **eggs**	18	0	9	0	579	620	10	10
blue eyes	**blue** tiger	drop **eyes**	21	0	6	0	546	711	5	11
dirty face	**dirty** hat	level **face**	13	0	6	0	594	778	9	2
right foot	**right** dog	choose **foot**	17	0	7	0	514	578	2	8
our house	**our** apple	sick **house**	36	0	12	0	1,399	485	3	10

Mean	*23* *(19 per mill.)*	*0*	*8*	*0*	*776*	*663*	*8* *(N = 60)*	*8* *(N = 61)*

Table 3. Prime-target bigram frequency and target word frequency for each maternal transcript.

Prime-Target	Prime-Target frequency, Target frequency															
	Anne	Aran	Becky	Carl	Dominic	Eleanor	Fraser	Gail	Joel	John	Lara	Liz	Nic	Ruth	Thomas	Warren
big boy	1,	18,	0,	13,	2,	1,	0,	1,	8,	0,	0,	0,	0,	1,	10,	12,
	26	190	16	470	56	21	229	18	78	18	11	14	10	31	205	127
in	17,	5,	13,	3,	11,	1,	8,	11,	5,	1,	6,	3,	9,	1,	2,	9,
bed	134	55	40	36	62	18	200	94	30	15	79	17	50	69	56	27
poor dolly	4,	30,	1,	0,	6,	0,	0,	3,	7,	6,	0,	2,	2,	0,	1,	5,
	98	278	38	15	60	1	0	27	46	115	10	38	55	22	1	19
some milk	23,	4,	10,	5,	2,	4,	4,	5,	11,	1,	3,	4,	1,	18,	19,	0,
	97	34	44	19	7	28	23	17	54	30	18	30	25	70	69	33
that box	17,	26,	5,	16,	8,	2,	7,	9,	5,	2,	4,	3,	1,	2,	7,	6,
	131	115	74	95	80	26	51	60	51	203	57	47	48	104	205	181
your hand	11,	42,	16,	18,	40,	24,	8,	14,	13,	13,	18,	13,	51,	11,	18,	19,
	28	69	33	54	71	45	20	27	26	45	45	24	92	30	33	58
good girl	34,	2,	54,	1,	1,	338,	0,	16,	1,	0,	107,	32,	175,	312,	1,	0,
	88	32	89	10	17	526	25	57	11	15	184	46	243	441	36	8
little baby	9,	0,	1,	15,	1,	4,	6,	0,	4,	8,	3,	2,	7,	22,	0,	7,
	475	69	164	183	18	41	107	73	75	62	98	64	133	333	92	217
new book	2,	3,	1,	1,	0,	1,	0,	1,	2,	3,	0,	2,	1,	0,	1,	1,
	76	82	18	54	31	67	93	66	132	76	72	84	87	123	78	65
next door	4,	0,	0,	0,	0,	0,	3,	5,	3,	1,	5,	0,	3,	0,	11,	0,
	50	93	42	50	33	19	54	57	40	44	37	32	41	94	100	35
pretty dress	4,	5,	0,	0,	0,	1,	0,	0,	0,	1,	1,	0,	0,	5,	0,	4,
	67	78	43	7	23	13	20	32	3	64	12	16	51	68	5	45
four eggs	1,	0,	1,	1,	2,	0,	0,	0,	1,	2,	0,	2,	4,	0,	0,	4,
	47	97	64	49	38	9	4	20	30	35	3	14	48	43	8	70
blue eyes	0,	0,	1,	1,	0,	0,	6,	0,	0,	10,	2,	0,	1,	0,	0,	0,
	45	41	26	25	7	49	55	18	17	51	20	38	17	69	26	42
dirty face	3,	0,	0,	3,	2,	0,	1,	0,	1,	0,	0,	0,	3,	0,	0,	0,
	31	24	43	51	34	17	7	34	44	18	41	17	61	53	53	66
right foot	0,	0,	0,	0,	0,	2,	1,	0,	3,	0,	1,	1,	3,	0,	0,	6,
	35	33	13	33	52	31	37	34	36	18	19	36	12	27	31	67
our house	8,	2,	0,	0,	0,	3,	1,	5,	0,	2,	1,	1,	3,	8,	1,	1,

147	189	63	80	36	39	83	76	44	76	68	50	50	243	124	31

Our experimental paradigm, illustrated in Figure 1, involves speaking aloud a word bigram that is either attested in child-directed maternal transcripts (206-frequency and 19-frequency experimental conditions) or not (control condition) while two pictures, the target and a distractor, are visually displayed. If the word bigram is familiar to children, then the first word in the bigram (the prime) may facilitate processing of the second word in the bigram (the target) such that the picture relating to the target is fixated more quickly than when word bigrams are unrelated.

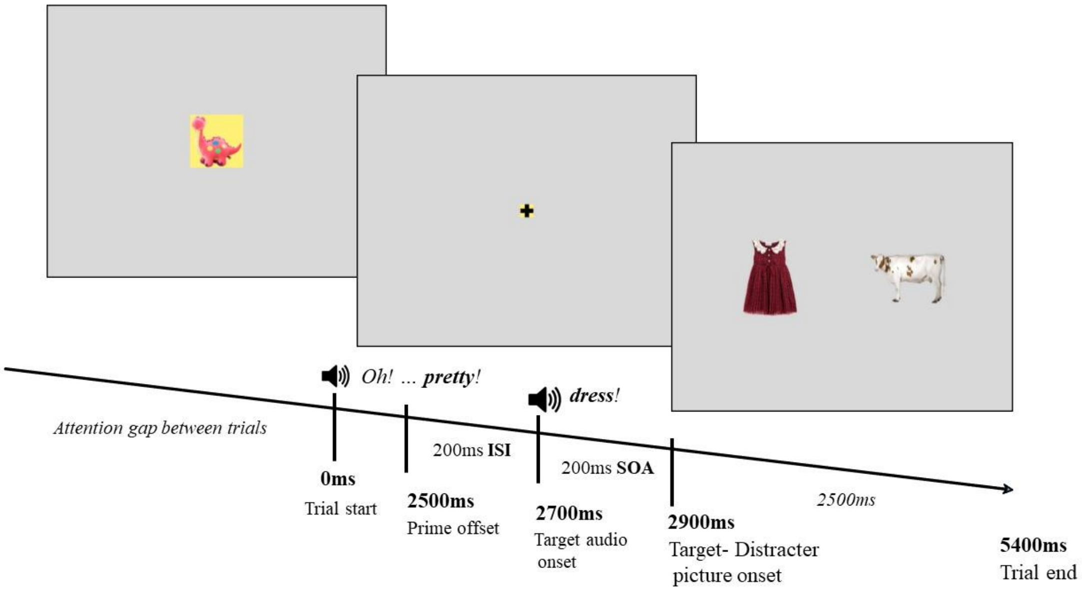

Figure 1. Example trial sequence for one experimental condition, with time in milliseconds given from the beginning of a trial. A control trial would consist of the same target but matched with a prime such that the prime-target word bigram was zero frequency (e.g., *safe dress*).

Procedure

Ethics approval was granted by the College Research Ethics Committee at Nottingham Trent University. Caregivers completed a consent form and a checklist to indicate those words in the current study that their child knew, and were handed a vocabulary inventory to assess vocabulary size (British communicative development inventory, Hamilton, Plunkett, & Schafer, 2000, completed by caregivers in their own time and returned by post). Vocabulary size results are not presented in the current paper due to low numbers of carers returning the inventory. Children were seated either in a car seat securely strapped to a chair or on their caregiver's knee in front of a 24-inch LCD monitor with an arm mounted Eyelink 1000 eyetracker set to a 500Hz sample rate. The monitor display was set at 1,280x1,024 pixels. The arm mount allowed the monitor and eye-tracker to be positioned approximately 60cm in front of the child's field of view. Auditory stimuli were delivered using loudspeakers placed on a table under the monitor. Due to attentional constraints on children of this age, there were only 16 trials in total per child; eight experimental (four 206-frequency, four 19-frequency) and eight control. The control stimuli used the remaining complement of targets after removing the experimental stimuli e.g., if the 206-frequency targets were *boy, bed, dolly,* and *milk* then the 206-frequency control targets would be *box, hand, girl,* and *baby*. Trials were counterbalanced across children such that within and between conditions, the target picture appeared to the left and right an equal number of times. The format of each trial can be seen in Figure 1. No participant was exposed to more than two consecutive trials within the same condition.

Results

The control condition grouped all control word bigrams together since all control bigrams are zero frequency (i.e., novel). Data manipulation and analysis were carried out using the R programming language (version 3.5.2; R Core Team, 2018), RStudio IDE for R (version 1.1.463; RStudio Team, 2016), lme4 (version 1.1.21, Bates, Mächler, Bolker, & Walker, 2015), nlme (version 3.1-140 ; Pinheiro, Bates, DebRoy, & Sarkar, 2019), lsmeans (Length, 2016), and nlmeU (version 0.70-3; Galecki & Burzykowski, 2013).

Despite making every effort to use primes and targets that children would know, 386 of 736 trials were discarded due to caregivers indicating their child did not know either the prime or target word for experimental and/or control conditions. A further 71 trials were discarded due to the child not looking at the target for the whole trial. This left 279 trials for analysis (number of trials [and infants]: 78 [33], 74 [35], and 127 [39] for the 206-frequency, 19-frequency, and control conditions respectively). For studies involving infants and children of similar ages to the current study, a sample size of over 30 is considered relatively large (see Oakes, 2017). Details of number of trials per word bigram are given in Table 2. Due to the number of trials that were discarded, for information purposes we provide post hoc power analyses (from Galecki & Burzykowski, 2013) for the key statistical results (see Table 4).

Table 4. Post-hoc power analysis for reaction time (raw and distractor-initial) with frequency and age as predictors.

	numDF	denDF	F-value	nc	Power
Raw reaction time					
Intercept	1	221	552.24	552.24	1.00
Frequency	2	221	2.57	5.14	.51
Age	1	221	.82	.82	.15
Frequency * Age	2	221	4.35	8.70	.75
Distractor-initial RTs					
Intercept	1	95	747.13	747.13	1.00
Frequency	2	95	6.84	13.68	.91
Age	1	33	8.23	8.23	.80

The area of interest for fixations was the 300x300 area of each picture plus an extra 60 pixels on all sides to account for any calibration error. We first present an overall picture of where children fixated during picture presentation by showing proportion of time spent looking at targets and distractors before focusing on the key research question regarding the speed by which children fixated on the targets.

Proportion of time spent looking at the target

The amount of time children spent looking at the target (T) and distractor (D) within each condition was used to calculate the proportion of time that children spent looking at the target (T / (T + D)). We include fixation data from the onset of the target and distractor pictures (i.e., after 2,900ms on trial) until the end of the trial (i.e., after 5,400 ms). Figure 2 shows the proportion of time spent looking at the target within each condition. As one would hope, children look at the target more often than the distractor in both experimental and control conditions (proportion of time looking at target > .50). Tables 5 and 6 show that there were no statistical differences across conditions in the proportion of time spent looking at the target. However, our central hypotheses concern processing efficiency, and so our key area of interest is whether there are any reaction time differences between experimental and control conditions. We present these analyses in the next section.

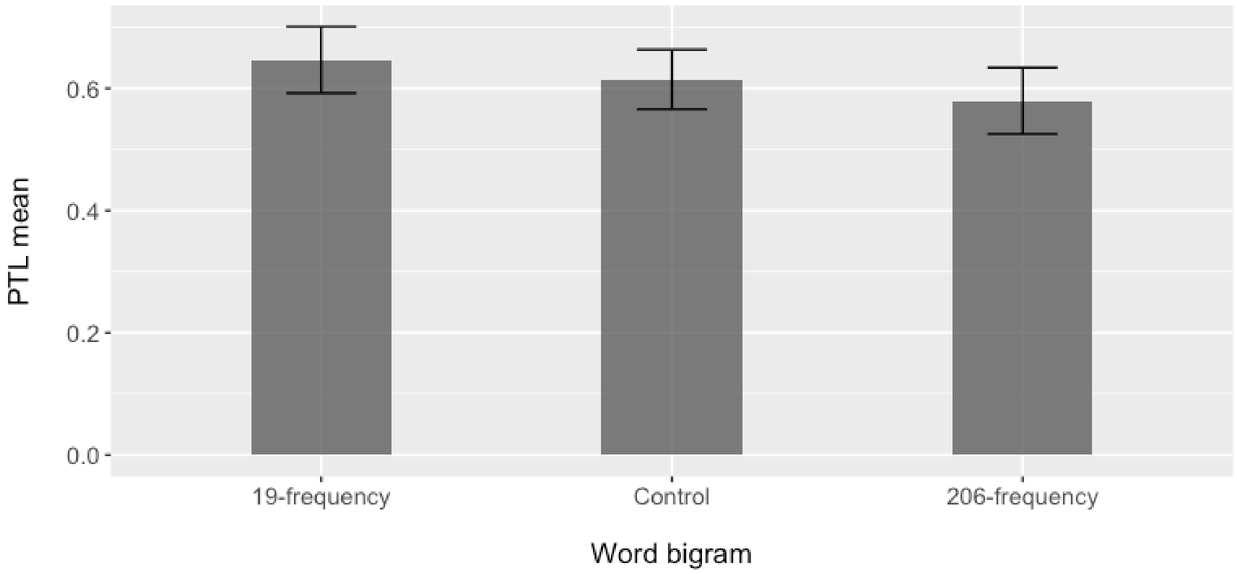

Figure 2. Mean proportion of time spent looking at the target relative to the distractor (PTL mean) within each condition, including 95% confidence intervals around the mean.

Table 5. Logistic mixed-effects model comparing proportion of time spent looking (PTL) at the target across conditions, using the 206-frequency condition as the baseline.

	PTL model		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	2.14	1.09 – 4.19	.028
206-freq vs. 19-freq	1.65	.67 – 4.10	.278
206-freq vs. Control	1.69	.76 – 3.78	.202
Random Effects			
σ^2	9.13		
τ_{00} Trial	5.84		
τ_{00} Item	.36		
τ_{00} Participant	.47		
ICC	.08		
N Trial	279		
N Item	62		
N Participant	45		
Observations	279		
Marginal R^2 / Conditional R^2	.01 / .09		

Table 6. Pairwise comparisons across experimental conditions and the control condition for the proportion of time spent looking at the target. Tukey was used to control for familywise error rates and Kenward-Roger to calculate degrees of freedom.

Contrast	Log estimate	SE	df	z.ratio	*p*
206-frequency - 19-frequency	-.50	.46	Inf	-1.08	.524
206-frequency - Control	-.52	.41	Inf	-1.28	.409
19-frequency - Control	-.02	.42	Inf	-.05	.998

Reaction time analyses

For these data, fixations are considered from 233ms after the onset of the target and distractor images (i.e., from 2900 + 233ms into the trial until trial end at 5,400 ms) to account for children mobilizing an eye movement (see Swingley, Pinto, & Fernald, 1999). Reaction times above 2 standard deviations were excluded (see Arnon & Snider, 2010). This left 227 trials for analysis (number of trials [and children]: 63 [32], 60 [29], and 104 [38] for the 206-frequency, 19-frequency, and control conditions respectively). The number of trials per stimulus is given in Table 2.

Reaction time to fixate on target (all trials)

Figure 3 shows reaction times in milliseconds to fixate on targets for children across the age range for each of the 19-frequency, 206-frequency, and control conditions. Table 7 shows the linear-mixed effects model comparisons using reaction time as the dependent variable and condition and age as independent variables. While condition has a marginal effect on reaction times, there is a clear interaction between condition and age, with reaction time reducing with age for the 19-frequency and 206-frequency word bigrams but not for the control word bigrams. Table 8 (upper table) explores the interaction model further by contrasting the control condition with the 206-frequency and 19-frequency conditions together with their interactions with age, with all contrasts being significant. To investigate whether there is any difference between experimental conditions, we contrasted the 206-frequency condition with the 19-frequency and control conditions (see Table 8, lower table). The only differences seen are those between experimental conditions and the control condition.

These results show that with age, children fixated on targets more quickly for the experimental conditions than for the control condition. All target response times are for word bigrams where both prime and target are known by the child and where primes and targets were audio-recorded as individual words. The difference between the experimental and control conditions therefore arises because in the experimental conditions the frequency of encounter of the word bigram will increase with age in child-directed speech whereas this is not the case for control word bigrams. There is no difference between experimental conditions, suggesting that the frequency by which the word bigram has appeared in child-directed speech has little bearing on the processing of the bigram, at least for bigrams beyond an average of 19 occurrences per million. For the above analysis, all trials were included regardless of the location of the initial fixation because of lack of convergence of the statistical model when considering only trials where the initial fixation was on the cross in the centre of the screen. We therefore support the above analyses by considering reaction times for those trials where the first image fixated on is the distractor (see Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998; Swingley, Pinto, & Fernald, 1999 for similar analyses).

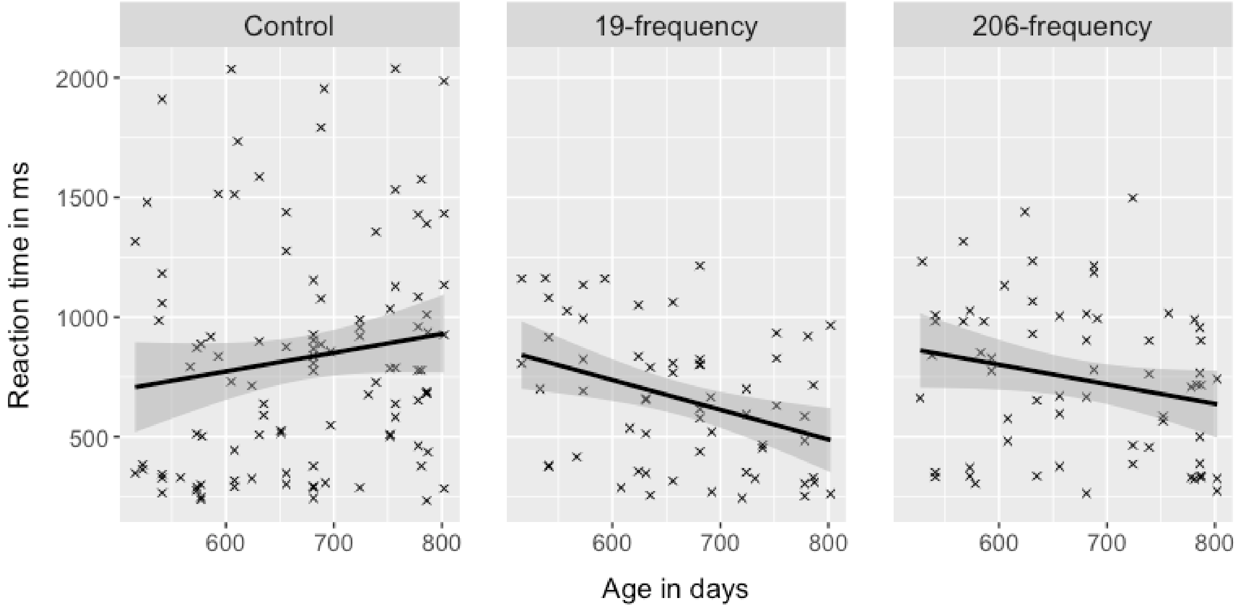

Figure 3. Reaction time (in milliseconds) to orient to the target image, with regression line and 95% confidence interval shadow.

Table 7. Linear mixed-effects model comparisons with reaction time as the dependent variable, and age and condition as the independent variables. Due to overfitting, for each model the random slope term was dropped and only the participant and item random intercepts included.

Model comparison	df	AIC	BIC	logLik	Deviance	Chisq	Chi Df	Pr(>Chisq)
Intercept model	4	3361	3375	-1677	3353	NA	NA	NA
Intercept vs. Condition	6	3361	3381	-1674	3349	5	2	.083
Intercept vs. Condition + Age	7	3362	3386	-1674	3348	6	3	.126
Intercept vs. Condition * Age	9	3357	3388	-1669	3339	15	5	.012

Table 8. Reaction time contrasts for the Condition * Age interaction model using either the control condition as baseline (upper table) or the 206-frequency condition as baseline (lower table), with the remaining conditions together with their interaction with age as the contrasts.

Raw RT model (control as baseline)

Predictors	*Estimates*	*CI*	*P*
(Intercept)	421.54	-167.28 – 1010.37	.161
Control vs. 206-frequency	909.06	7.11 – 1811.02	.048
Control vs. 19-frequency	1087.00	146.19 – 2027.81	.024
Age	.60	-.26 – 1.47	.172
Control vs. 206-frequency * Age	-1.48	-2.80 – -.16	.028
Control vs. 19-frequency * Age	-1.87	-3.26 – -.48	.008
Random Effects			
σ^2	129280.33		
τ_{00} Item	16238.68		
τ_{00} Participant	4608.87		
ICC	.14		
N Item	62		
N Participant	44		
Observations	227		
Marginal R^2 / Conditional R^2	.07 / .20		

Raw RT model (206-frequency as baseline)

Predictors	*Estimates*	*CI*	*p*
(Intercept)	1330.61	600.95 – 2060.26	<.001
206-frequency vs. Control	-909.06	-1811.02 – -7.11	.048
206-frequency vs. 19-frequency	177.93	-857.47 – 1213.34	.736

Age	-.88	-1.94 – .19	.106
206-frequency vs. Control * Age	1.48	.16 – 2.80	.028
206-frequency vs. 19-frequency * Age	-.39	-1.92 – 1.14	.616
Random Effects			
σ^2		129280.33	
$\tau_{00\ \text{Item}}$		16238.68	
$\tau_{00\ \text{Participant}}$		4608.87	
ICC		.14	
N_{Item}		62	
$N_{\text{Participant}}$		44	
Observations		227	
Marginal R^2 / Conditional R^2		.07 / .20	

Reaction time to fixate on target for distractor-initial trials

For this analysis, only trials where the distractor was the first of the two images fixated on are included (132 trials). Figure 4 shows the reaction times and Table 9 the linear mixed-effects model comparisons with reaction time as the dependent variable and condition and age as the independent variables. There was an effect of both condition and age on reaction time but no interaction between the two. We therefore conducted pairwise comparisons across conditions (see Table 10), showing that children in both experimental conditions fixated on the target more quickly than the control condition, with no difference between experimental conditions. The analyses are broadly similar to those above: fixations to target are faster in the experimental conditions than the control condition, with no difference between experimental conditions.

While our analyses show a clear effect between control and experimental conditions on the speed to look at the target words contained in word bigrams, they are not in line with our expectation that looking times would be faster for the 206-frequency condition than the 19-frequency condition. We shall return to this in the Discussion.

Processing of two-word sequences

However, there are some caveats to the results of Study 1. First, our stimuli had to be constrained to ensure that both experimental and control primes appeared in caregiver utterances, control prime-target pairs did not occur in caregiver utterances, and all targets occurred a substantial number of times. As a result, some of the control word bigrams are potentially less meaningful as word pairs than the experimental word bigrams (e.g., *start dolly* vs. *poor dolly*) while some control primes are also verbs (e.g., *start*) that may inhibit processing of the target (e.g., Thorpe & Fernald, 2006). Second, our analyses required the removal of trials due to parents reporting their child did not know either the prime and/or the target (the reduced sample size could also potentially explain the lack of any difference between experimental conditions). Third, while Study 1 suggests processing efficiencies for word bigrams appearing 19 or more times per million in caregiver utterances, it does not highlight whether such efficiencies may influence subsequent word learning. We therefore supplement our findings with a direct analysis of corpus data in Study 2 to: (a) examine only word bigrams that have been produced in caregiver utterances; (b) examine a fuller range of word bigram frequencies; and (c) examine whether the frequency of the attested word bigrams influences the learning of nouns appearing immediately after the bigram.

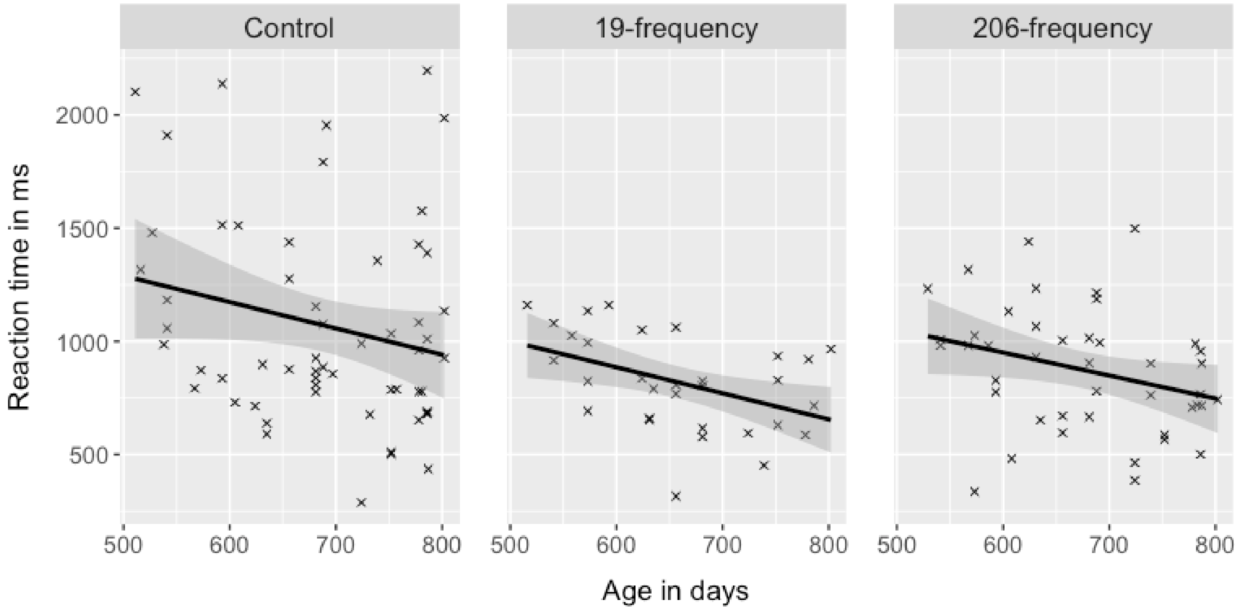

Figure 4. Reaction time (in milliseconds) to orient to the target image when the initial fixation was on the distractor image, with regression line and 95% confidence interval shadow.

Table 9. Linear mixed-effects model comparisons for children who first fixated on the distractor rather than the target, with reaction time to fixate on the target as the dependent variable, and age and condition as the independent variables. Due to overfitting, the random slope term and item random intercept were dropped and only the participant random effect included. Contrasts are for the Condition + Age model.

Model comparison	df	AIC	BIC	logLik	Deviance	Chisq	Chi Df	Pr(>Chisq)
Intercept model	3	1938	1947	-966	1932	NA	NA	NA
Intercept vs. Condition	5	1930	1944	-960	1920	13	2	.002
Condition vs. Condition + Age	6	1923	1941	-956	1911	8	1	.004
Condition + Age vs. Condition * Age	8	1927	1951	-956	1911	0	2	.982

Predictors	*Estimates*	*CI*	*p*
(Intercept)	1875.74	1321.69 – 2429.78	<.001
Control vs. 206-frequency	-206.19	-338.60 – -73.77	.002
Control vs. 19-frequency	-264.07	-412.57 – -115.56	<.001
Age	-1.18	-1.98 – -.37	.004
Random Effects			
σ^2	109069.00		
τ_{00} Participant	10015.20		
ICC	.08		
N Participant	35		
Observations	132		
Marginal R^2/ Conditional R^2	.15 / .22		

Table 10. Pairwise comparisons between conditions for reaction time to fixate on target when the first fixation for the infant was on the distractor. Tukey was used to control for familywise error rates and Kenward-Roger to calculate degrees of freedom.

Contrast	Estimate	SE	df	t.ratio	*p*
Control – 206-frequency	206.19	67.90	116	3.04	.008
Control – 19-frequency	264.07	76.43	120	3.46	.002
206-frequency – 19-frequency	57.88	80.26	121	0.72	.751

Study 2 – Corpus analyses of maternal word n-grams and child vocabulary acquisition

Study 1 indicates that children process two-word sequences occurring an average of 19 times per million and 206 times per million in caregiver speech more quickly than controls (i.e., novel two-word sequences), with no difference between levels of frequency. If children are able to quickly process relatively frequent two-word sequences in an utterance, then this may provide more resources to focus on the novel aspects of the utterance (e.g., novel words). As such, children may learn more novel words within utterances containing relatively frequent multi-word sequences than those that do not. We test this hypothesis in Study 2.

Method

The same corpus as Study 1 was used, with each mother-child transcript being examined individually. Although we are primarily interested in word bigrams, we also investigated the effect of word trigram and word four-gram frequencies on subsequent word learning to see whether there were any differential effects depending on the length of the multi-word sequence.

For each mother-child transcript, we extracted all nouns produced by the mother and used the child utterances to determine those maternal nouns that were also *subsequently* produced by the child (the *learned* set) and those that the child did not produce (the *not learned* set). Note that the learned set does not include nouns that were produced by the child before being produced by the mother so that we only assess nouns that are plausibly learned. Since our analysis involves secondary corpus data from mother-child transcriptions, child noun production is the existence of the noun within the transcribed child utterances (i.e., the transcriber perceived the noun as having been produced, but we do not know whether that production was correct on a phoneme-by-phoneme basis or not). From the resulting sets of nouns, we then removed any that the mother produced in isolation, since these may be easier to learn than when the noun is within a multi-word utterance (e.g., Swingley & Humphrey, 2018).

Since the learned and not learned sets of nouns could differ on key word-level predictors of word learning, we directly controlled for their frequency in caregiver speech and their phonemic length by: (a) removing high frequency nouns from the learned set until the learned and not learned sets were matched for their frequency of use in caregiver speech; and

(b) removing long nouns from the not learned set until the learned and not learned sets were matched for phonemic length. Relevant descriptive statistics for pre- and post-matching are given in Table 11. As we will see, other word-level factors that influence word learning, such as concreteness and neighborhood density, are controlled for indirectly by including them as predictors in the regression analysis.

Table 11. Mean *N*, log frequency and phonemic length of nouns across maternal utterances in mother-child transcripts for the learned and not learned sets of nouns, before and after frequency and length matching of the sets.

N-gram	Set	Pre-match			Post-match		
		N	Log10(freq)	Length	*N*	Log10(freq)	Length
Bigram	Not learned	102	.79	4.30	98	.78	4.15
	Learned	138	2.21	3.96	37	.78	4.20
Trigram	Not learned	99	.81	4.30	96	.80	4.14
	Learned	138	2.22	3.96	36	.80	4.19
Four-gram	Not learned	93	.85	4.30	88	.84	4.07
	Learned	134	2.27	3.94	33	.83	4.15

From the resulting set of learned and not learned nouns, we extracted the word bigrams, trigrams and four-grams that existed in all maternal utterances that contained the noun. For each noun, we then calculated the average frequency of word bigrams, trigrams and four-grams based on their frequencies within maternal utterances across the whole transcript. For example, if ball was in the learned noun set and only appeared in the five-word utterance "*Which one is the ball?*", the mean frequency for word bigrams would be the average frequency of *which one*, *one is,* and *is the*; the mean frequency for word trigrams would be the average of *which one is* and *one is the,* and the four-gram frequency would be the frequency of *which one is the*. Based on this method, for each mother-child transcript, we could see whether learned nouns appeared in utterances having a higher average word bigram/trigram/four-gram frequency than not learned nouns.

Results

Since we use maternal word bigram/trigram/four-gram frequency as a predictor of child word learning, frequency is log10(frequency) (see Brysbaert, Mandera & Keuleers, 2018). Note that frequency refers to the caregiver utterances i.e., the nouns under consideration for the children are matched on their frequency of occurrence in caregiver utterances.

Figure 5 graphically compares the mean bigram/trigram/four-gram frequency of utterances involving learned and not learned nouns. Table 12 provides the corresponding statistical analyses, with three mixed-effects logistic regression models (bigram, trigram, and four-gram) predicting word learning as a function of log10 multi-word frequency. In these analyses we also include key word-level factors as predictors of word learning: neighborhood density, phonotactic probability and concreteness. Although the values for these three variables are the same regardless of whether we examine bigram, trigram, or four-gram frequency, we include them as predictors in all analyses for thoroughness. We also standardise all predictors to reach models convergence. The analyses show that concreteness is a significant predictor of word learning with neighborhood density and phonotactic probability having little effect. However, crucially, the analyses also show that learned nouns appear in utterances having significantly higher bigram and trigram frequencies, but there is no difference in four-gram frequency between the learned and not learned nouns. As the familiarity of the word bigrams and word trigrams within utterances increases, so does the likelihood of children learning the nouns contained in those utterances.

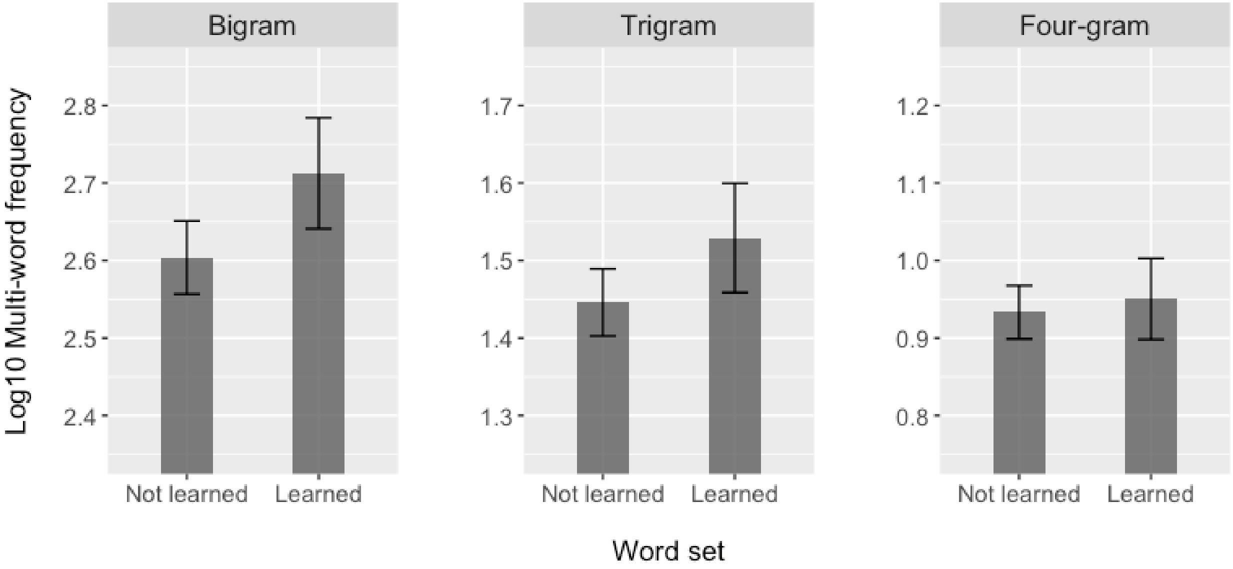

Figure 5. Comparison of the mean log10(multi-word frequency) of utterances in caregiver speech that contain learned nouns versus nouns that were not learned (nouns being matched

for caregiver word frequency and for phonemic length). Frequencies are the average frequency of either all bigrams, trigrams, or four-grams within caregiver utterances that contained either a learned or not learned noun.

Table 12. Mixed-effects logistic models for bigram, trigram and four-gram sets. The dependent variable is word learned (0, 1). The independent variables are log10(multi-word sequence frequency), neighborhood density, phonotactic probability, and concreteness. Child was used as a random effect.

	Bigram model		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	.33	.25 – .45	**<.001**
Bigram frequency	1.15	1.04 – 1.28	**.007**
Neighborhood density	1.06	.96 – 1.18	.238
Phonotactic probability	1.03	.93 – 1.14	.597
Concreteness	2.01	1.75 – 2.30	**<.001**
Random Effects			
σ^2	3.29		
$\tau_{00\ \text{child}}$	.31		
ICC	.09		
N_{child}	16		
Observations	2159		
Marginal R^2 / Conditional R^2	.13 / .20		

	Trigram model		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	.34	.25 – .46	**<.001**
Trigram frequency	1.15	1.04 – 1.27	**.007**
Neighborhood density	1.06	.95 – 1.18	.278
Phonotactic probability	1.02	.92 – 1.13	.708
Concreteness	2.01	1.75 – 2.31	**<.001**
Random effects			
σ^2	3.29		

$\tau_{00\ \text{child}}$	.31
ICC	.09
N $_{\text{child}}$	16
Observations	2108
Marginal R^2 / Conditional R^2	.13 / .20

	Four-gram model		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	.34	.25 – .46	**<.001**
Four-gram frequency	1.06	.95 – 1.18	.293
Neighborhood density	1.08	.97 – 1.20	.164
Phonotactic probability	1.02	.92 – 1.14	.678
Concreteness	1.96	1.70 – 2.26	**<.001**
Random effects			
σ^2	3.29		
$\tau_{00\ \text{child}}$	.34		
ICC	.09		
N $_{\text{child}}$	16		
Observations	1924		
Marginal R^2 / Conditional R^2	.11 / .20		
N $_{\text{Child}}$	16		
Observations	6248		
Marginal R^2 / Conditional R^2	.00 / .08		

Discussion

In two studies we have shown that (1) two-year-old children process word bigrams more efficiently when they have appeared in child-directed maternal transcripts than when they have not, and (2) the frequency of word bigrams in maternal utterances relates to children's production of nouns that appear in those utterances. These results have major implications for early language development by suggesting that two-year-old children not only have knowledge of some of the word bigrams appearing in their linguistic input but also these word bigrams may subsequently be processed more efficiently, giving greater opportunity to focus on other aspects of the linguistic stimulus.

Processing of two-word sequences

While Study 1 found that hearing word bigrams having frequencies of 19+ per million resulted in faster looking times to the noun target within the bigram, there was no difference between bigram frequencies averaging 19 times per million and bigrams averaging 206 per million. This may well be because bigrams with frequencies of 19 per million or greater occupy the most frequent 5% of word bigrams. There are only 7,519 of 162,339 word bigrams having a mean frequency of 19+ occurrences per million. In essence, our experimental conditions both reflect high frequency bigrams. Unfortunately it is somewhat difficult to use bigrams of lower frequencies while still having noun targets that would be expected to be known to young children. The real issue though is the number of utterances that contain a familiar word bigram. Of the 401,308 utterances in the corpora used, 248,323 (61.9%) contain three or more words (i.e., utterances where the presence of a word bigram could affect processing of a word that is not part of the bigram) and 247,470 of those 248,323 (99.7%) contain a word bigram with a frequency of 19 or more per million. In a nutshell, almost two thirds of all utterances that young children hear contain a word bigram that is sufficiently familiar to them to facilitate their processing of the remaining words in the utterance.

The findings from Study 1 add to the existing literature that has indirectly examined word bigrams. In this work, highly familiar bigrams have been used. For example, Fernald, Pinto, Swingley, Weinberg and McRoberts (1998) found that with age, 15- to 24-month old children move their eyes more quickly to pictures of nouns spoken at the end of a familiar carrier phrase, with final word bigram frequencies averaging 131 per million. Bannard and Matthews (2008) found final word bigram frequency to be a significant predictor of the time to speak aloud the first three words in their four-word phrases, with the lowest word bigram frequency being around 2,568 per million. What we show here is that these efficiencies apply to word sequences below the phrase level and also apply to word bigrams that are far lower in frequency than previous work suggests.

When we considered the fuller range of word bigram frequencies (Study 2), we found that the likelihood of children producing a noun was related to the frequency of the word bigrams within the caregiver utterances that carried the noun. This supports previous work that suggests linguistic processing efficiencies increase vocabulary uptake (e.g., Lany, 2018; Mahr & Edwards, 2018). Moreover, the frequency by which a word bigram is heard greatly depends on the language to which the child is exposed. This potentially provides an explanation for larger vocabularies in children who hear more caregiver speech: these children are more able to learn new vocabulary because they can process word bigrams within utterances more quickly based on having heard the word bigrams more often.

The idea that language is based on usage is not new (e.g., Bybee, 2006; Ellis, 2002) and here we evidence that simple word pairs are processed more efficiently so long as they have been encountered a small number of times in natural language. In many ways, this should not come as a surprise: 8-month-old infants extract some knowledge of sequential transitions across syllables after a 2-minute exposure (Saffran, Aslin & Newport, 1996), so the same mechanism is likely to be at work at the larger grain size of the word-level. Yet, the suggestion that familiarity with word bigrams affects processing of the remainder of an utterance once again brings to bear issues surrounding quantity and diversity in this age group. Vocabulary acquisition seems best suited to a language input that is initially rich in quantity and subsequently rich in lexical diversity (Jones & Rowland, 2017; Rowe, 2012). Quantity is important for word bigrams to become sufficiently familiar to the child such that efficiencies can be gained to process other material in the utterance. Later diversity is necessary because further increases in exposure do not make processing any more efficient (see Study 1) – instead, efficiencies are achieved by having a larger bank of word bigrams to exploit. If this is the case, then it follows that those children who enjoy a linguistic input that is rich in diversity will also have a learning experience that affords greater opportunity to learn more new words, and so the rich get richer (see also Weisleder & Fernald, 2013).

It is also worth noting that in our study, children over ~21 months fixate on a picture of a noun within a familiar word bigram some 200-300 ms more quickly than nouns in a control (i.e., zero frequency) condition. Contrast this with adult priming effects of less than 100 ms even when the same word is used as both prime and target (Ostergaard, 1998) or around 30 ms for collocates that share similarities to some of our word bigrams (Durrant & Doherty, 2010). The magnitude of our effects is consistent with other priming studies using young children (see Arias-Trejo and Plunkett, 2013, for semantic priming). In their case and ours, a relatively abstract experimental paradigm was used to investigate prime-target effects, but there is no reason to suspect that processing advantages within natural language discourse would be greatly dissimilar to those seen here.

While our results indicate that two-year-old children’s language processing may be facilitated by familiar two-word sequences, there are aspects of our studies that need to be borne in mind. For example, in Study 1 we were unable to fully control the two-word sequence stimuli so that they were identical in form (e.g., control stimulus matched to experimental stimulus for word category) and so we supplemented our findings with a corpus-based study. In Study 2 (the corpus study) we had to make a necessary assumption that children who failed to produce a noun had ‘not learned’ the noun, whereas of course it may be the case that noun production was not captured within the particular transcripts used. In addition, there could be particular features of the nouns produced by children in the learned set that are different to those of the not learned set, such as being more child-friendly. Unfortunately, no database exists containing all possible predictors of word learning for entry into statistical analyses. That said, a word’s frequency in caregiver speech is the most potent predictor of word learning (e.g., Hansen, 2016; Swingley & Humphrey, 2018) and the nouns under consideration in the corpus study were frequency-matched, giving some assurance that word learning may be facilitated by the frequency of two-word sequences.

In summary, two-year-old children process word bigrams occurring 19 or more times per million in caregiver speech more quickly than novel word bigrams. Subsequent corpus analyses show a positive relationship between children’s noun production and the frequency of the word bigrams that appear in the utterances containing the noun, over and above the frequency of the noun itself. In combination, these findings suggest that not only do children have knowledge of two-word sequences at a very young age but also that the processing of the remainder of the utterance is facilitated, allowing learning to focus on the more novel aspects of an utterance.

XXX developed the study concept. XXX and ZZZ contributed to the study design. Testing and data collection were performed by YYY and ZZZ. YYY performed the data analysis. All authors approved the final version of the manuscript for submission.

Neither of the studies reported in this article was formally preregistered. Requests for the data can be sent via email to the lead author.

Acknowledgements

ZZZ was supported by a PhD scholarship from the German Catholic Academic Exchange Service and the Leibniz Association. We are grateful to Petya Ventsislavova for her support during this work and to Mark Torrance and Jens Roeser for their suggestions and technical support during the experimental setup. We thank all families who participated in the study.

References

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15,* 802–814.

Arias-Trejo, N., & Plunkett, K. (2013). What's in a link: Associative and taxonomic priming effects in the infant lexicon. *Cognition*, *128,* 214–227.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, *62,* 67–82.

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science, 19,* 241–248.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48.

Bleses, D., Makransky, G., Dale, P. S., Højen, A., & Ari, B. A. (2016). Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics*, *37*, 1461–1476.

Bybee, J. (1998). The emergent lexicon. *Chicago Linguistic Society, 34*, 421–435.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language, 82*, 711–733.

Case, R. (1978). Intellectual development from birth to adulthood: A neo-Piagetian interpretation. In *Children's thinking: What develops?*, Chi, M. T. H., & Siegler, R. (Eds.), pp. 37–71.

Durrant, P., & Doherty, A. (2010). Are high-frequency collocations psychologically real? Investigating the thesis of collocational priming. *Corpus Linguistics and Linguistic Theory*, *6*, 125–155.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*, 143–188.

Fernald, A., & Hurtado, N. (2006). Names in frames: Infants interpret words in sentence frames faster than words in isolation. *Developmental Science, 9*, F33–F40.

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

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, *72,* 1003–1015.

Gałecki, A., & Burzykowski, T. (2013). *Linear mixed-effects models using R: A step-by-step approach*. New York: Springer.

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language*, *27,* 689–705.

Hansen, P. (2017). What makes a word easy to acquire? The effects of word class, frequency, imageability and phonological neighbourhood density on lexical development. *First Language*, *37,* 205–225.

Howes, D. H., & Solomon, R. L. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology, 41*, 401–410.

Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science, 11*, 31–39.

Jones, G., Cabiddu, F., Andrews, M., & Rowland, C. F. (2020). Phonological knowledge and its substantive role in children's early word learning. Under review, *Developmental Science*.

Jones, G., & Macken, B. (2015). Questioning short-term memory and its measurement: Why digit span measures long-term associative learning. *Cognition*, *144,* 1–13.

Jones, G., & Macken, B. (2018). Long-term associative learning predicts verbal short-term memory performance. *Memory and Cognition*, *46,* 216–229.

Jones, G., & Rowland, C. F. (2017). Diversity not quantity in caregiver speech: Using computational modeling to isolate the effects of the quantity and the diversity of the input on vocabulary growth. *Cognitive Psychology*, *98,* 1–21.

Jones, G., Tamburelli, M., Watson, S. E., Gobet, F., & Pine, J. M. (2010). Lexicality and frequency in specific language impairment: Accuracy and error data from two nonword repetition tests. *Journal of Speech Language and Hearing Research*, *53,* 1642–1655.

Keren-Portnoy, T., Vihman, M. M., Depaolis, R. A., Whitaker, C. J., & Williams, N. M. (2010). The role of vocal practice working memory. *Journal of Speech, Language, and Hearing Research*, *53,* 1280–1293.

Lany, J. (2018). Lexical-processing efficiency leverages novel word learning in infants and toddlers. *Developmental Science, 21,* 1–12, e12569.

Lenth, R., V. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software*, *69,* 1–33.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, *18,* 193–198.

Lieven, E., Salomo, D. & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics, 20*, 481–508.

MacWhinney, B. (2000). *The CHILDES project: The database* (Vol. 2). Psychology Press.

Mahr, T., & Edwards, J. (2018). Using language input and lexical processing to predict vocabulary size. *Developmental Science, 21,* e12685.

Munson, B., Kurtz, B. A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 48,* 1033–1047.

Oakes, L. M. (2017). Sample size, statistical power, and false conclusions in infant looking-time research. *Infancy, 22,* 436–469.

Ojima, S., Matsuba-Kurita, H., Nakamura, N., & Hagiwara, H. (2011). The acceleration of spoken-word processing in children's native-language acquisition: An ERP cohort study. *Neuropsychologia, 49,* 790–799.

Ostergaard, A. (1998). The effects on priming of word frequency, number of repetitions, and delay depend on the magnitude of priming. *Memory & Cognition, 26,* 40–60.

Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D. (2019). *nlme: Linear and nonlinear mixed effects models*. R package version 3.1-143.

R Core Team (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech vocabulary development. *Child Development, 83,* 1762–1774.

Rowland, C. F. (2007). Explaining errors in children's questions. *Cognition, 104,* 106–134.

Rowland, C. F. & Fletcher, S. L. (2006). The effect of sampling on estimates of lexical specificity and error rates. *Journal of Child Language, 33,* 859–877.

RStudio Team (2016). *RStudio: Integrated development for R*. RStudio, Inc., Boston, MA. http://www.rstudio.com/

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928.

Swingley, D. (2007). Lexical exposure and word-form encoding in 1.5-year-olds. *Developmental Psychology, 43,* 454–464.

Swingley, D., & Humphrey, C. (2018). Quantitative linguistic predictors of infants' learning of specific English words. *Child Development, 89,* 1247–1267.

Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71*, 73–108.

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language, 28*, 127–152.

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology, 67*, 1176–1190.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science, 24*, 2143–2152.

Zamuner, T. S. (2009). Phonotactic probabilities at the onset of language development: Speech production and word position. *Journal of Speech, Language, and Hearing Research, 52,* 49–60.