University of Amsterdam

UvA-DARE (Digital Academic Repository)

CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Determinants of early lexical acquisition: Effects of word- and child-level factors on Dutch children's acquisition of words

Josje VERHAGEN[1], Mees VAN STIPHOUT[1], and Elma BLOM[2]

[1]Amsterdam Center for Language and Communication University of Amsterdam, the Netherlands and [2]Department of Special Education: Cognitive and Motor Disabilities Utrecht University, The Netherlands
Address for correspondence: Josje Verhagen, P.O. Box 1637, 1000 BP Amsterdam, The Netherlands, j.verhagen@uva.nl

Abstract

Previous research on the effects of word-level factors on lexical acquisition has shown that frequency and concreteness are most important. Here, we investigate CDI data from 1,030 Dutch children, collected with the short form of the Dutch CDI, to address (i) how word-level factors predict lexical acquisition, once child-level factors are controlled, (ii) whether effects of these word-level factors vary with word class and age, and (iii) whether any interactions with age are due to differences in receptive vocabulary. Mixed-effects regressions yielded effects of frequency and concreteness, but not of word class and phonological factors (e.g., word length, neighborhood density). The effect of frequency was stronger for nouns than predicates. The effects of frequency and concreteness decreased with age, and were not explained by differences in vocabulary knowledge. These findings extend earlier results to Dutch, and indicate that effects of age are not due to increases in vocabulary knowledge.

Introduction

Previous studies on the factors that contribute to early lexical acquisition indicate that children acquire frequent and concrete words before infrequent and abstract words (Bates, Dale & Thal, 1995; Braginsky, Yurovsky, Marchman & Frank, 2019; Hansen, 2017), and phonologically simple words before more complex ones (Storkel, 2004; Vihman & Croft, 2007). Factors associated with the age of acquisition of words are often correlated: for example, frequent words tend to be shorter (Zipf, 1936) and more concrete (Reilly & Kean, 2007) than less frequent words. Hence, it is crucial that multiple factors are included in one analysis to assess the independent and relative contribution each factor makes to acquisition. Only few studies to date

CrossMark

have taken such a multi-factor approach (Braginsky et al., 2019; Hansen, 2017; Storkel, 2009; Swingley & Humphrey, 2018). In this earlier work, however, acquisition data were pooled across children, which leaves unaddressed whether differences between individual children in age, gender, and family background impacted on the outcomes. In the current study, we complement earlier research in three ways. First, we investigate which word-level factors predict children's acquisition of words in a new language: Dutch. Second, by including not only word-level factors, but also child-level factors (i.e., age, gender, parental education), we assess the relative contributions of word-level factors, once variation at the child level is controlled. Finally, we assess whether the effects of the word-level factors interact with age, and if so, whether these interactions can be attributed to increases in children's vocabulary knowledge.

Word-level properties that predict early lexical development

Of all factors that may shape children's early lexical development, frequency has received most attention. Ample evidence indicates that words with high input frequency are acquired before less frequent words (e.g., Ambridge, Kidd, Rowland & Theakston, 2015; Goodman, Dale & Li, 2008). Analyzing mother-child interactions, for example, Huttenlocher, Haight, Bryk, Seltzer, and Lyons (1991) found that the order in which English 14- to 26-month-olds acquired new words was predicted by the frequency with which their mothers used these words. Goodman et al. (2008) correlated word frequencies in child-directed speech with acquisition data from the American MacArthur Bates Communicative Development Inventory (CDI) (Fenson, Marchman, Thal, Dale, Reznick & Bates, 2007) and found that words that were frequent in child-directed speech were acquired earlier.

Frequency is not the sole factor driving children's early acquisition of words, however. Semantic factors such as concreteness or imageability and phonological factors, such as neighborhood density, word length and phonological complexity, play a role as well. Concreteness refers to the degree to which a concept denoted by a word refers to a perceptible entity (Brysbaert, Stevens, De Deyne, Voorspoels & Storms, 2014) (e.g., 'chair' is more concrete than 'love'). Children acquire more concrete words before less concrete words (Bird, Franklin & Howard, 2001; Braginsky et al., 2019). Concreteness is closely tied to imageability, or the ease with which a word provokes a mental image or sensory experience (Bird et al., 2001). In fact, high correlations have been found between concreteness and imageability (> .8, Richardson, 1976), and imageability – like concreteness – correlates with age of lexical acquisition. For example, McDonough, Song, Hirsh-Pasek, Golinkoff, and Lannon (2012) found that the degree to which words were rated as imageable correlated with children's age of acquisition based on CDI data, above and beyond word class (see Ma, Golinkoff, Hirsh-Pasek, McDonough & Tardif, 2009 for Chinese-speaking children). Concreteness also relates to word class such that nouns (at least the ones in children's surrounding input) are generally more concrete than verbs, which, in turn, are more concrete than adverbs, adjectives, and prepositions. This has given rise to the noun dominance hypothesis (Gentner & Boroditsky, 2001) according to which concrete nouns are acquired first because nouns – unlike verbs and other word classes – typically label perceptually salient referents about which children already have concepts, through their experiences with the world.

As for phonological factors, earlier work has shown that shorter words predominate in children's early vocabularies, rather than longer words (Storkel, 2004, 2009), and

words with no or few consonant clusters are acquired before words with such clusters (Storkel, 2004; Vihman & Croft, 2007). Furthermore, neighborhood density, or the number of words that can be formed by adding or removing one phoneme of a word or by replacing one phoneme with a different phoneme, has been found to affect early word production. Specifically, words from dense phonological neighborhoods developmentally precede words from sparser neighborhoods (Coady & Aslin, 2003; Stokes, 2010; Storkel & Lee, 2011), although this effect might be subject to individual variation (Storkel & Maekawa, 2005) or disappear once frequency is controlled (Swingley & Humphrey, 2018).

This last observation demonstrates that word properties can be correlated. Frequent words, for example, are generally shorter than less frequent words – a finding that has become known as Zipf's law (Zipf, 1936) and holds true across languages (Balasubrahmanyan & Naranan, 2008; Ferrer-i-Cancho, 2005). Frequent words are also often phonologically simpler, due to phonological reduction (Bybee, 2010). As for word length, correlations with other word properties indicate that shorter words tend to have more phonological neighbors than longer words (Pisoni, Nusbaum, Luce & Slowiaczek, 1985; Storkel, 2004) and are typically more concrete (Reilly & Kean, 2007). Word class may also interfere with these factors, such that nouns are generally less frequent (Goodman et al., 2008) and more concrete (Simonsen, Lind, Hansen, Holm & Mevik, 2013) than verbs. Due to these correlations, it is crucial that multiple factors are taken into account when examining determinants of children's early word productions: Such research enables not only a more valid assessment of the effects, but also allows an investigation of the relative importance of each factor.

Multi-factor studies on early lexical acquisition

Only a handful of studies have investigated the simultaneous impact of multiple word-level factors on young children's lexical development (Braginsky et al., 2019; Hansen, 2017; Storkel, 2009; Swingley & Humphrey, 2018). Swingley and Humphrey (2018) investigated the effects of frequency, concreteness, word duration, neighborhood density and phonotactic probability on lexical acquisition in eight English-speaking 12- to 15-month-olds, and found that frequency (based on mothers' speech) showed the largest effects, followed by concreteness. Word length approached significance, and neighborhood density and phonotactic probability were not significant. Globally similar results were obtained by Hansen (2017), who examined the effects of frequency, concreteness, word class, and phonological neighborhood density on acquisition data from 5,674 Norwegian children aged between 8 and 32 months. Following common approaches (Goodman et al., 2008; Ma et al., 2009; McDonough et al., 2012), Hansen determined the age of acquisition for each word: that is, the age at which a certain threshold – in this case, at least 50% of the children – produced this word. The author found that word frequency (based on child-directed speech) was the most important predictor of age of acquisition, followed by imageability and word length. Neighborhood density did not predict children's acquisition. Furthermore, nouns were acquired before verbs, but this effect was mediated by the effect of imageability. Finally, in a large-scale study by Braginsky and colleagues (2019) using the CDI, acquisition data from over 38,000 children learning 10 different languages were investigated to examine the effects of frequency, concreteness, arousal, valence, word length, mean length of utterance (i.e.,

the mean length of the utterance in which a word occurred in child-directed speech), and babiness (i.e., the degree to which a word is associated with babies). Braginsky and colleagues not only assessed the independent contribution of each factor, but also whether the effects of these factors interacted with age and lexical category (i.e., nouns, predicates, function words). The results showed that, across languages, frequency, concreteness, mean length of utterance, and babiness were stronger predictors of age of acquisition than word length, valency, and arousal. Effects of frequency and concreteness became stronger with age (i.e., from about 8 to 30–36 months), such that the tendency to acquire words that are frequent and concrete words as opposed to less frequent and more abstract was stronger in older than in younger children, and the effect of frequency was stronger for nouns than function words. Since most of these findings held true across languages, the authors concluded that the same principles guide early word learning cross-linguistically.

Taken together, earlier multi-factor studies converge on the finding that frequency and concreteness (or imageability) are most important and that phonological properties such as neighborhood density and word length play a less important role (Braginsky et al., 2019; Hansen, 2017; Swingley & Humphrey, 2018). They also indicate that effects of these factors may interact with age and word class, such that effects of frequency and concreteness increase with age (Braginsky et al., 2019) and effects of frequency are stronger for nouns than function words (Braginsky et al., 2019; Hansen, 2017).

Two issues remain unaddressed in these earlier studies, however. First, it is as yet unclear to what extent the results are influenced by child-level factors. Across studies, acquisition data were aggregated over children, such that the dependent variable in the analysis was either 'age of acquisition' of each word (i.e., the age at which a certain threshold – usually 50% – of the children have acquired the word, cf. Hansen, 2017) or the number of children comprehending/producing each word (Braginsky et al., 2019). The drawback of aggregating data over children is that individual child characteristics such as gender and socio-economic background cannot be taken into account. Earlier work has shown that girls are generally better word learners than boys (Eriksson et al., 2012; Huttenlocher et al., 1991) and children from higher socio-economic backgrounds generally know more words than children from lower socio-economic backgrounds (Hart & Risley, 1995; Rowe & Goldin-Meadow, 2009). For socio-economic status, moreover, there is some indirect evidence suggesting that differences may not only be quantitative but also qualitative in nature, such that children from lower socio-economic backgrounds generally are exposed to less diverse and less sophisticated vocabulary (Huttenlocher , Waterfall, Vasilyeva, Vevea, & Hedges, 2010) and less decontextualized input (Rowe, 2012), which, in turn, are related to vocabulary development (Rowe, 2012).

Second, it is currently unclear how the interactions between word-level factors and age should be interpreted. In Braginsky et al. (2019), effects of frequency and concreteness increased with age. A possible explanation of these results is that children's reliance on these word-level factors became stronger as their vocabularies increased. Earlier research has shown that effects of factors involved in early word learning may be modulated by vocabulary growth. Phonological memory, for example, has been found to be more important at early stages of acquisition, when children's vocabularies are small, than at later stages (Verhagen, Boom, Mulder, de Bree & Leseman, 2019), while other principles, including syntactic bootstrapping (Moyle, Weismer, Evans & Lindstrom, 2007) and mutual exclusivity (Lewis,

Cristiano, Lake, Kwan & Frank, 2020), have been found to become more important as children's vocabularies increase. These findings suggest that the strength of the effects of factors involved in word learning may vary as a function of children's vocabulary knowledge. To what extent increases in vocabulary knowledge account for the earlier-attested interactions between age and word-level factors (i.e., frequency and concreteness) is as yet an open question.

The current study

In this study, we examine how word frequency, concreteness and phonological factors (neighborhood density, word length, number of consonant clusters) relate to early lexical acquisition in Dutch, taking into account inter-individual differences in age, gender, and parental education (as a proxy for socio-economic background). We also investigate how any effects of such factors may interact with age and word class, and – in case interactions with age are found – whether these are due to increases in children's vocabulary knowledge. Specifically, our research questions were the following:

1) How do frequency, concreteness, word class, and phonological factors predict the acquisition of words in Dutch-speaking toddlers, once differences in age, gender and parental education are controlled?
2) Do any effects of the word-level factors vary with word class and age?
3) If interactions between the word-level factors and age are found, can these be attributed to differences in children's level of vocabulary knowledge?

With respect to the first question, we expected that more frequent and more concrete words would be acquired before less frequent and less concrete ones, respectively (Braginsky et al., 2019; Hansen, 2017; Swingley & Humphrey, 2018). Furthermore, we expected that nouns would be acquired before other word classes (Gentner & Boroditsky, 2001). Finally, we predicted that words from more dense neighborhoods, shorter words and words with no or few consonant clusters would be acquired before words from sparser neighborhoods, longer words and words with many consonant clusters, but with effects much smaller in magnitude than those of frequency and concreteness (Hansen, 2017; Swingley & Humphrey, 2018). Regarding the second question, we predicted that any effect of concreteness would be stronger for nouns than for other word classes (Braginsky et al., 2019). We also predicted that if effects of frequency and concreteness were found, these would increase with age (Braginsky et al., 2019). Finally, concerning the third question, we had no a priori predictions. Thus, the question of whether interactions with age would be accounted for by differences in children's level of vocabulary knowledge was exploratory.

Method

Participants

Data were taken from a longitudinal study in the Netherlands (pre-COOL, cf. Mulder, Verhagen, van der Ven, Slot & Leseman, 2017; Verhagen et al., 2019) in which children's linguistic and cognitive development was followed between two and five years of age. For the present study, data collected at the first measurement were used, when children were between 22 and 42 months old. Children's parents had filled out

the short form of the Dutch version of the MacArthur Bates Communicative Developmental Inventory (i.e,. N-CDI short form of "Words and Sentences" version B, Zink & Lejaegere, 2002, henceforth referred to as CDI), as well as a questionnaire assessing child and family characteristics.

The original sample of subjects contained 1,523 participants. From this sample, children were excluded if they (i) were from homes in which another language than Dutch was spoken next to or instead of Dutch ($n = 397$) or (ii) could be considered late talkers using the commonly used criterion of scoring below the 10% percentile on the CDI for their age and gender ($n = 51$). The latter group was excluded since children with expressive language delays have been shown to have a weaker noun bias than children without such delays (Jiménez, Haebig & Hills, 2020; MacRoy-Higgins, Shafer & Fahey, 2016). Moreover, stronger effects of neighborhood density have been observed for late talkers (Stokes, 2010, 2014), as well as a weaker preference for words that are semantically related to words already acquired: that is, a stronger preference for "oddball words" (Beckage, Smith & Hills, 2011). These differences indicate that the lexicons of children with low expressive vocabularies for their age are not only quantitatively but also qualitatively different from those of more typical talkers. An additional 45 children were not included either, due to missing data for age, gender, or parental education level. This resulted in a sample of 1,030 children who ranged between 22 and 42 months in age when their parents filled out the CDI ($M = 28.76$ months, $SD = 3.61$ months), 528 of which were boys (47.26%). The CDI is intended for children aged up until 30 months (for norm data and more information about reliability and validity of the Dutch short form of the CDI, see Zink & Lejaegere, 2003). However, given substantial variation in scores of children between 30 and 42 months in our sample, and post-hoc analyses that showed that our results did not change as a function of whether the older children were included, children above 30 months were included. Children's parents' educational level was generally high. Specifically, parents' reports of their highest educational level in a questionnaire using a 4-point scale ranging from 1 ('primary education') to 4 ('higher education') showed a mean value of 3.36 ($SD = 0.68$, ratings averaged over parents).

Data

The short form of the Dutch CDI (N-CDI, Zink & Lejaegere, 2002) consists of a list of 116 words and short phrases from various categories, such as foods, animals, and toys. The list is a subset of the long form of the Dutch CDI that is a close translation of the American CDI "Words and Sentences" (Fenson et al., 1994). For each word or phrase, parents indicate whether their child 'understands' or 'understands and says' the word or phrase. In the current study, only expressive data were recorded (i.e., "Does your child understand and say this word?"), because the passive component of the CDI has been shown to be unreliable for children older than 16 months (Bates et al., 1995).

In the current study, not all 112 items could be included. First, items were excluded if they involved semantically related words that are given as alternatives in the N-CDI, such as *warm/heet* 'warm/hot' ($n = 25$). Such alternatives typically differ in frequency, word length, consonant clusters, and phonological neighborhood density, making it impossible to assign values for such word-level factors to these items. Second, closed-class words such as prepositions and interjections were excluded ($n = 9$), as these form a small and diverse set of words, including pronouns, interjections,

prepositions, particles, and question words. Third, onomatopoeia such as *beh beh* 'baa [sound sheep]' (*n* = 8) were excluded, due to lack of frequency and concreteness ratings for these items. Finally, multi-word utterances such as *ik wil* 'I want' (*n* = 5) were excluded because values for some of the word-level factors (e.g., concreteness and neighborhood density), could not be calculated for these items. The resultant set of items contained 65 items, including 45 nouns and 20 predicates (10 verbs, 8 adjectives, 2 adverbs) (see Appendix A for the complete list).

Word-Level Factors
Word class
Each item of the CDI was categorized as either a Noun or a Predicate. Further distinctions within the class of predicates (i.e., between verbs and adjectives) were not made, due to the relatively small number of items in each class.

Frequency
Input frequency of each item was determined based on token frequencies from the parent and interviewer tiers in three Dutch CHILDES corpora (Bol & Kuiken, 1990; van Kampen, 2009; Wijnen & Elbers, 1993). Token frequencies were used, because, in Dutch, derivations can be highly phonologically distinct from their lemmas. For example, for the word *koe* ('cow'), rather distinct forms are used for the diminutive from *koetje* ('little cow') or the plural *koeien* ('cows'). Frequencies were log-transformed for the analyses.

Concreteness
For each item in the present study, concreteness values were obtained from previously collected concreteness ratings for 30,000 Dutch words, gathered from Dutch adult native speakers through a 5-point scale ranging from 1 'very abstract/language-based' to 5 'very concrete/experience-based' (Brysbaert et al., 2014). These ratings have been shown to be highly reliable (intra-class correlation, .92, Brysbaert et al., 2014), and strongly correlated with imageability ratings from another corpus (.73) (Brysbaert et al., 2014; van Loon-Vervoorn, 1985).

Neighborhood density
Neighborhood density, or the number of words that can be formed by adding or removing one phoneme or replacing one phoneme with a different phoneme (e.g., pear / fair / hair / heir) was calculated for each item with the help of the *Test Neighbors* software for Dutch (Pasveer, 2004), using the software PhonoTactools (Adriaans, 2006). The CELEX corpus was used as input in this program, a 42-million-word corpus of written texts (Baayen, Piepenbrock & van Rijn, 1993). Neighborhood density was missing for one item (*frietjes* 'fries'), as this item did not occur in the CELEX database. Values were log-transformed for the analyses.

Word length
For each item, the number of phonemes was counted. Diphthongs were counted as a single phoneme. For forms ending in /ən/ (e.g., *lopen* 'walk' and *buiten* 'outside'), the /n/ was not included in the phoneme count, as it is usually dropped by Dutch speakers (i.e., *lopen* 'walk' is pronounced as /loːpə/, cf. Booij, 1999).

Number of consonant clusters

For each item, the number of consonant clusters was counted. CC-clusters as well as CCC-clusters were counted as one cluster. A binary variable was then created, representing words with versus without consonant clusters.

Child-level factors

Age of the child in months at the time parents filled in the CDI was recorded. The child's gender was recorded by the parents as either 'boy' or 'girl'. Parents' highest completed level of education was reported on a 4-point scale having 1 ('primary education'), 2 ('vocational training') 3 ('secondary education'), and 4 ('higher education'), as its scale points. In two-parent households, the mean value of education of the two parents was taken as a measure of the family's education level. Vocabulary was assessed with an independent measure; the Dutch version of the Peabody Picture Vocabulary Test (PPVT-III-NL, Dunn & Dunn, 2005). The PPVT is a receptive vocabulary test in which children choose one out of four pictures after an orally presented word. To reduce testing time, a shortened version was used in the current study that contained a fixed set of 24 items. This shortened version was obtained by removing items that did not differentiate well across children because they were either too easy or too difficult (for more details on the properties of this test, see Verhagen et al., 2019). Internal consistency of the test was very good (Cronbach's alpha = .85). Scores were calculated as percentages correct for each child, and age-residualized scores were used for the analyses, because children's age at time of testing differed from their age at the time of their parents' filling out the CDI. Specifically, age at time of testing (in months) was regressed on children's sum scores on the PPVT and the resultant, residualized scores were used for analysis, so as to obtain a measure of children's receptive vocabulary scores in which the variance due to age was taken out (for a similar procedure, see Mulder et al., 2017). Vocabulary scores were available for 998 children (97%).

Data screening and analyses

Prior to our main analyses, multicollinearity between the word-level predictor variables was checked through Pearson correlations for all continuous predictor variables and a point-biserial correlation for the binary predictor variable word class. Subsequently, we fitted three generalized linear mixed-effects models to address our research questions. First, to address our question on the effects of the word- and child-level factors on children's acquisition of the CDI words, we fitted a mixed-effects model with *Acquired* as the dependent variable, a binary variable (0 or 1), representing whether an item had been acquired by the child. Three fixed-effect child-level factors were included: the continuous variable *Age* (number of months), the binary variable *Gender* (boy or girl), and the continuous variable *Parental Education* (ranging between 1.0 and 4.0). Five word-level factors were included: the binary variable *Word class* (Noun or Predicate), the continuous variable *Frequency* (frequency values based on CHILDES, ranging between 1 and 1052, log-transformed for the analyses), the continuous variable *Concreteness* (concreteness ratings from Brysbaert et al. (2014), ranging between 1.47 and 5.00), the continuous variable *Neighborhood density* (number of neighbors, ranging between 0 and 36, log-transformed for the analyses), the continuous variable *Word length* (number of phonemes, ranging between 2 and 12),

and the binary variable *Consonant clusters* (recoded as 'No clusters' or 'Clusters'). To address our second question, we fitted the same model, with additional two-way interaction terms between the fixed-effect factors that were significant in the previous model and *Age,* and between these factors and *Word class.* Finally, to address our third question, we ran a final model in which two additional interactions were added between the continuous *Vocabulary* (PPVT score) and the word-level factors that were significant in the second model. The generalized linear mixed-effects models were fitted in R (R Core Team, 2014), using the *lme4* package (Bates, Maechler, Bolker & Walker, 2015). In all models, by-item and by-subject random intercepts were included, as well as a by-item random slope for *Age* and *Vocabulary* (in models two and three). Orthogonal sum-to-zero contrasts were set for the binary variables *Gender, Word Class, and Consonant clusters.* All continuous variables were centered around the mean. The *glmer* function supplies significance indicators such as *p*-values from zero, and a log odds estimate measure of effect size. However, since log odds values can be difficult to interpret, they were exponentiated to odds for ease of interpretation. In our results section, odds values are reported; for a full summary of the model that includes the log odds values, see Appendix B. By default, the *bobyqa* optimizer was used to fit the models. If a model failed to converge, the *allFit* function was employed to find a more suitable optimizer, as recommended by Bates and colleagues (2015). Explained variance of the model was computed using the *2rglmm* package (Jaeger, 2017), based on R^2 for generalized linear mixed-effects models from Johnson (2014). Our scripts and data files are available as supplemental materials in the OSF databased, at https:// osf.io/n84mk/?view_only=3270384a02a2450ca25b04bf072a7770.

Results

Descriptive statistics and correlations

On average, children were reported to have acquired 49 out of 65 items on the Dutch CDI ($SD = 15$, min-max = 0–65). Descriptive statistics for the word- and child-level factors are shown in Table 1 (for the child-level variable *Gender*, see "Participants"); correlations between the continuous word-level predictor variables are shown in Table 2.

The highest correlation was between concreteness and word class ($r(65) = -.74$, $p < .001$), which indicated that nouns had higher concreteness ratings than predicates. A moderate to high negative correlation was found between word length and neighborhood density, indicating that shorter words generally had more neighbors than longer words ($r(64) = -.63$, $p < .001$). The remaining correlations were weak to moderate (rs between .01 and .33).

Effects of the word-level and child-level predictors on acquisition

A generalized linear regression model on the data for each word (1 = acquired, 0 = not acquired) with all word- and child-level fixed-effect predictors showed that the intercept point estimate of the probability that a word was acquired was 2.69 in log odds, which means that, on average, the probability of a word being acquired was 0.93 (95% CI: 0.92 .. 0.94, $p < .001$).[1] This baseline predicts the acquisition of an item for a theoretical

---

[1]Note that the intercept is not indicative of the actual average probability that a word is acquired by the child: rather, it is the probability of acquisition of a hypothetical item with all the word-level factors set to

Table 1. Descriptive Statistics for the Word-Level and Child-Level Factors

	M	SD	Min	max	N
Word-level factors					
Frequency	97.42	197.23	1	1052	65
Concreteness	4.26	0.95	1.47	5.00	65
Word length	5.14	1.99	2	12	65
Consonant clusters	0.54	0.58	0	2	65
Neighborhood density	9.16	8.50	0	36	64
Child-level factors					
Age	28.76	3.61	22	42	1030
Parental education	3.34	0.68	1	4	1030
Vocabulary	66.09	18.77	0	100	998

Note. Non-transformed data are given here, but note that log-transformed data for frequency and neighborhood density, and age-residualized vocabulary scores were used in the analyses.

Table 2. Bivariate Correlations Among the Word-Level Predictors

	Frequency	Concreteness	Word length	Consonant clusters
Concreteness	−.30*	–	–	–
Word length	−.22	.04	–	–
Consonant clusters	−.26*	.03	.31**	–
Neighborhood density	.16	.09	−.63***	−.28*
Word class	.33**	−.74***	−.01	.01

Note. *** $p < .001$, ** $p < .01$, * $p < .05$

situation in which the subject is a boy of 28.76 months old (the mean age) and whose parents have an education score of 3.34 (the mean child-level factor values) and the baseline item has the mean values for all word-level predictors (see Table 1).

Regarding the word-level predictors, the model showed two main effects. First, there was a positive effect of frequency: on average, a word was 7.23 times more likely to be acquired with every increase in frequency by one standard deviation: that is, 197 occurrences in the child-directed speech corpus (estimate (odds): 7.228, 95% CI: 5.625 .. 9.289, $p < .001$). Second, there was a positive effect of concreteness: on average, a word was 2.44 times more likely to be acquired with every increase in concreteness by one standard deviation, or 0.95 points (estimate (odds): 2.443, 95%

---

the average by a hypothetical child with all the subject-level factors set to the average. When age is not mean-centered, but set to the lowest value (i.e., 22 months), the model yields a much lower intercept probability that a word is acquired of .02. Importantly, however, irrespective of whether factors are mean-centered, the estimates of the effects remain exactly the same.

CI: 1.945 .. 3.068, $p < .001$). The word-level predictors word class, word length, consonant clusters, and phonological neighborhood density were not significant.

Regarding the child-level factors, two effects were found. First, there was a positive effect of age: on average, a word was 1.25 times more likely to be acquired with every increase in age by one standard deviation: that is, 3.61 months (estimate (odds): 1.253, 95% CI: 1.229 .. 1.277, $p < .001$). Second, there was a positive effect of gender: on average, a word was 1.90 times more likely to be acquired if the child was a girl (estimate (odds): 1.901, 95% CI: 1.663 .. 2.174, $p < .001$). Parental education was not a significant predictor of a word's acquisition (estimate: (odds): 1.176, 95% CI: 1.063 .. 1.300, $p = .107$). For a full overview of the results, see Table B1 in Appendix B.

Interactions with age and word class

To assess our second question, a mixed-effects model was run in which interactions were added between age and the factors that came out significant in the previous model – frequency and concreteness – as well as the interactions between these factors and word class. The interaction word class*concreteness could not be included as this yielded non-converging models, presumably due to the high correlation between these variables ($r = .74$, see Table 2). In this new model with interaction terms, the same main effects were found as in the first model: that is, there were positive effects of age (estimate (odds): 1.233, 95% CI: 1.208 .. 1.258, $p < .001$), gender (estimate (odds): 1.903, 95% CI: 1.665 .. 2.175, $p < .001$), frequency (estimate (odds): 6.953, 95% CI: 5.469 .. 8.841, $p < .001$), and concreteness (estimate (odds): 2.039, 95% CI: 1.616 .. 2.571, $p < .001$). In addition, the model yielded three significant interactions. First, there was a negative interaction between age and frequency, such that the effect of frequency in odds decreased by 0.043 with every increase in age by 3.61 months (estimate (odds): 0.957, 95% CI: 0.947 .. 0.968, $p < .001$). Second, we found a significant interaction between age and concreteness, which indicated that the effect of concreteness in odds decreased by 0.033 with every increase in age by 3.61 months (estimate (odds): 0.967, 95% CI: 0.960 .. 0.9738, $p < .001$). Third, the model yielded an interaction effect between word class and frequency such that the effect of frequency decreased by 0.65 for predicates, compared to nouns (estimate (odds): 0.355, 95% CI: 0.229 .. 0.551, $p = .018$). For the full model results, see Table B2 in Appendix B.

Disentangling the effects of age and vocabulary knowledge

To address our final question of whether the interactions between age and the word-level factors (i.e., frequency and concreteness) were due to increases in children's vocabulary, a final model was run. In this model, two additional interactions were entered: vocabulary*frequency and vocabulary*concreteness. Prior to running this model, the correlation between age and vocabulary (i.e., percentage correct scores) was computed, to check for multicollinearity. Vocabulary and age correlated moderately and positively ($r(997) = .35$, $p < .001$), indicating no multicollinearity. The results of this model in which both age and vocabulary were added as interaction terms with frequency and concreteness showed several effects. First, vocabulary had a positive effect, such that children who had higher scores on the receptive vocabulary test were more likely to have acquired a word on the CDI than children with lower scores: on average, a word was 1.889 times more likely to

be acquired with every increase in receptive vocabulary by one standard deviation (estimate (odds): 1.889, 95% CI: 1.7613 .. 2.026, $p < .001$). Second, the main effects and significant interactions as found in the previous model remained, including those for age. Specifically, this model showed that a word was 1.22 times more likely to be acquired with every increase in age by 3.61 months (estimate (odds): 1.223, 95% CI: 1.199 .. 1.248, $p < .001$). Third, none of the interactions with vocabulary size were significant. Thus, while age interacted with both frequency and concreteness in this model, no interactions were found between vocabulary and these factors.

The conditional R_c^2, representing how much variance was explained by the full model, was 0.703, 0.698, and 0.691 for the first, second, and third model respectively. The marginal R_m^2, representing how much variance can be explained by the fixed variables, was 0.207, 0.207, and 0.227, respectively. This indicates that, in all models, substantial variation was found at the level of children and items.

Discussion

This study set out to answer three questions: (i) Which word-level factors predict early word acquisition?, (ii) Do the effects of those factors vary with age and word class?, (iii) Can any interactions with age be attributed to individual differences in vocabulary knowledge? Specifically, we aimed to contribute to earlier work in which multiple word-level factors were included (Braginsky et al., 2019; Hansen, 2017; Swingley & Humphrey, 2018) by investigating a new language (Dutch), including not only word-, but also child-level factors, and disentangling effects of age and differences in level of vocabulary knowledge. The word-level factors in our analysis were word frequency (based on child-directed speech), concreteness, word class, word length, number of consonant clusters, and neighborhood density. The child-level factors were age, gender, and parental education. Acquisition data obtained through the short form of the Dutch CDI from over 1,000 Dutch toddlers were analyzed.

With regard to the first question, the data showed significant effects of frequency and concreteness, but not of the other word-level factors: word length, the presence of consonant clusters in a word, and neighborhood density. The effect of frequency is in keeping with earlier work showing that more frequent words are acquired before less frequent words (Braginsky et al., 2019; Goodman et al., 2008; Hansen, 2017; Swingley & Humphrey, 2018). In our study, the effect of frequency was substantial: while controlling for age and all other word-level predictors, a word that was one standard deviation more frequent than the mean was 7.23 times more likely to be acquired. This supports earlier conclusions that frequency is a predictor of acquisition, all else being equal (Ambridge et al., 2015) and is in line with the studies by Hansen (2017) and Braginsky et al. (2019) in which frequency came out as one of the strongest predictors.

Regarding concreteness, we found that concrete words were more likely to be acquired than more abstract words, in line with earlier studies (Bird et al., 2001; Braginsky et al., 2019; Hansen, 2017). Importantly, this effect was found even with word class included in the analysis, replicating earlier work that concreteness contributes to word learning within lexical categories (Braginsky et al., 2019; Swingley & Humphrey, 2018). It is important to note, however, that the mean concreteness value in our sample was high (i.e., mean value of 4.2 on a 5-point scale), presumably because of the large number of nouns in the Dutch CDI.

The effects of frequency and concreteness in our study remained when variance explained by other word-level factors as well as variation at the child-level due to

age, gender and parental education were controlled. At the child-level, effects of age and gender were found, but not of parental education. The gender effect is in line with earlier findings showing an advantage for girls over boys (Bornstein, Hahn & Haynes, 2004; Eriksson et al., 2012), and has been attributed to neurodevelopmental factors such as earlier brain lateralization in girls (Eriksson et al., 2012). The lack of an effect for parental education in our study contradicts earlier work showing effects of socio-economic status on children's vocabulary acquisition (Hart & Risley, 1995; Hoff, 2003) and might be due to limited variation and an overall high educational level in the current sample. Future work with more diverse samples is needed to confirm this.

None of the phonological factors in our study – word length, consonant clusters, or phonological neighborhood density – predicted acquisition. This was not unexpected, given the inconclusive findings in earlier work for these factors. For word length, for example, some studies showed an effect (Braginsky et al., 2019; Hansen, 2017; Storkel, 2004, 2009), whereas others did not (Swingley & Humphrey, 2018). Braginsky et al. (2019) found an effect of word length for function words only – a class that we did not include due to scarcity of items. Future research could establish the role of word length, including more fine-grained analyses involving several word classes, including function words. Regarding neighborhood density, one direction for future research would be to include familiarity ratings in neighborhood density scores such that only words that are rated as familiar to the child are included (Storkel, 2004) or calculate frequency-weighted neighborhood density values, taking into account the frequency of the neighbors (Vitevitch & Luce, 1999) to approximate more closely the type of language that children are exposed to.

Concerning our second question regarding possible interactions between the word-level factors and age and word class, we found – for word class – that the effect of frequency was stronger for nouns than predicates. This finding is somewhat similar to results in Braginsky et al. (2019) who found that frequency was one of the strongest predictors for nouns and predicates, but not for function words, although the word classes in our study differ from those studied in Braginsky et al. It is also in line with results by Goodman et al. (2008) who found that, within word classes (e.g., nouns, verbs), words that are more frequent in speech to children are likely to be learned earlier (Goodman et al., 2008). Our data demonstrated no main effect of word class. Earlier work showed that effects of word class disappeared when concreteness was included (Hansen, 2017) and suggested that variation in concreteness explains differences in acquisition between nouns and verbs (Gentner, 1982). This may hold true for the current data as well: there were no significant effects of word class, and concreteness was strongly correlated with word class. Note, however, that the number of word classes in our study was limited and that the class 'predicates' contained verbs, adjectives and adverbs, which may have made an effect of word class hard to detect.

As for interactions with age, we found that the effects of frequency and concreteness decreased with age, such that older children were less likely to be affected by frequency and concreteness in their acquisition than younger children. These findings contradict those of Braginsky and colleagues (2019) who found that the effects of frequency and concreteness INCREASED with age. However, a closer look at Braginsky et al.'s data shows considerable variation in the frequency*age and concreteness*age interactions across the languages investigated, with some languages showing positive interactions, but others showing very weak or even negative interactions. Since Braginsky et al. did not specify which languages showed positive or negative interactions with age, it is currently an open question how our results for Dutch relate to their results at the

level of individual languages. The finding that, in our data, the effects of frequency and concreteness decreased with age may be explained along the lines of the 'leveraged learning' account (McMurray, 2007; Mitchell & McMurray, 2009). This account posits that as children acquire more words, new words become easier to learn, through a set of processes, including fast mapping, mutual exclusivity, and syntactic bootstrapping, that become more efficient when children know more words. Specifically, the acquisition and knowledge of a 'starter set' of words positively impacts on the processes needed for word learning, and, in turn, on children's learning of subsequent words. Indeed, earlier work has shown that vocabulary knowledge positively predicts children's reliance on mutual exclusivity, and, in fact, is a better predictor than age (Lewis et al., 2020). Similarly, syntactic bootstrapping becomes more important as children's vocabularies increase, at least in typically-developing children (Moyle et al., 2007). Assuming that this 'leveraged learning' account holds true, one might predict that word-level properties such as frequency and concreteness play an increasingly less important role, as item-independent processes such as fast mapping, mutual exclusivity and syntactic bootstrapping become more important. This might then explain why, in our data, effects of frequency and concreteness decreased with age. However, on this 'leveraged learning' account, the interactions between age and the word-level factors would be expected to be driven by differences in vocabulary knowledge. This is not what we found. Rather, our data showed that the interactions between age and frequency/concreteness remained and no interactions between vocabulary and frequency/concreteness appeared, once vocabulary knowledge was added to our analysis. This could signal that factors associated with age, other than vocabulary knowledge, drove the effects. Alternatively, a possible explanation of why differences in vocabulary knowledge, unlike age, did not interact with the word-level factors is methodological in nature: our measure of individual differences in vocabulary knowledge involved a receptive vocabulary task (PPVT), while our outcome measure (CDI) reflected expressive knowledge. Although receptive and expressive vocabulary tend to be moderately to strongly correlated in young children (Bornstein & Hayes, 1998), the receptive task in our study may have tapped different processes or different aspects of word knowledge than the expressive (CDI) measure, which, in turn, may have attenuated any effects of vocabulary knowledge on word-level properties in our data. In addition, while the CDI assessed children's language abilities through parent reports, the PPVT involved a direct assessment of the child, which leaves open the possibility that the lack of interaction with vocabulary in our study was due to a difference in type of assessment. Future research could investigate in more detail whether differences in children's vocabulary knowledge modulate effects of age on word-level factors in early word learning, using better-aligned vocabulary measures that do not vary as to whether they assess receptive or expressive skills and are either based on parental report or direct assessments of children's language knowledge. Furthermore, studies could examine the associations between word-level factors and age/vocabulary by modelling these as non-linear, so as to obtain a more detailed picture of how effects may wax and wane across specific ages and depending on children's vocabulary size.

This study has a few limitations. First, since data were collected through the short form of the CDI and a number of items had to be excluded, the number of words analyzed was limited. Relatedly, we were unable to investigate function words, and collapsed verbs, adjectives and adverbs into one category called 'predicates'. We recommend that future work replicates the current study using the long version of the N-CDI, including function words as well as larger numbers of items per word

class, to test the robustness of the current effects of frequency and concreteness and examine in more detail how they relate to different word classes. A second limitation is that the children were mostly from families with highly-educated parents, which may have prevented us from finding a significant effect of parental education. Also, this limits the generalizability of our results.

To conclude, the current results support earlier research showing that frequency and concreteness are more important for children's acquisition of words than phonological factors in a new language (Dutch), and when variation in child-level factors (i.e., age, gender, parental education) is controlled. We also replicated the earlier finding that effects of frequency and concreteness vary with age (albeit in a different direction than in earlier work) and that the effect of frequency was stronger for nouns as opposed to predicates. Furthermore, we demonstrated that the origin of the effects of age does not lie in a factor highly correlated with age: increases in vocabulary knowledge. Taken together, our results provide further evidence that, across languages, the frequency with which words appear in the input and the degree to which words denote concrete things are among the most important determinants of early word learning, especially at early stages of acquisition.

References

Adriaans, F. (2006). PhonotacTools (Test version) [computer program]. Utrecht Institute of Linguistics OTS, Utrecht University, the Netherlands.

Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language*, *42*, 239-273. doi: 10.1017/S030500091400049X.

Baayen, H., Piepenbrock, R., & Van Rijn, H. (1993). The CELEX lexical database. Linguistic data consortium. Philadelphia, PA: University of Pennsylvania.

Balasubrahmanyan, V. K., & Naranan, S. (2008). Quantitative linguistics and complex system studies. *Journal of Quantitative Linguistics*, *3*, 177–228. doi: 10.1080/09296179608599629

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. doi: 10.18637/jss.v067.i01.

Bates, E., Dale, P. S., & Thal, D. (1995). Individual differences and their implications for theories of language development. *The Handbook of Child Language*, *30*, 96–151.

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

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers*, *33*, 73–79. doi: https://doi.org/10.3758/BF03195349.

Bol, G. W., & Kuiken, F. (1990). Grammatical analysis of developmental language disorders: A study of the morphosyntax of children with specific language disorders, with hearing impairment and with Down's syndrome, *Clinical Linguistics and Phonetics*, *4*, 77–86. doi: https://doi.org/10.3109/02699209008985472.

Booij, G. (1999). *The phonology of Dutch*. Oxford: Oxford University Press.

Bornstein, M. H., & Haynes, O. M. (1998). Vocabulary competence in early childhood: Measurement, latent construct, and predictive validity. *Child Development*, *69*, 654–671. doi: 10.1111/j.1467-8624.1998.tb06235.x

Bornstein, M., Hahn, C.-S., & Haynes, O. (2004). Specific and general language performance across early childhood: Stability and gender considerations. *First Language*, *24*(3), 267–304. doi: 10.1177/0142723704045681.

Braginsky, M., Yurovsky, D., Marchman, V. A., & Frank, M. C. (2019). Consistency and variability in children's word learning across languages. *Open Mind: Discoveries in Cognitive Science*, 3, 52–67. doi: 10.1162/opmi_a_00026.

Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica*, 150, 80–84. doi: 10.1016/j.actpsy.2014.04.010.

Bybee, J. (2010). *Language, usage and cognition* (Vol. 98): Cambridge University Press Cambridge.

Coady, J. A., & Aslin, R. N. (2003). Phonological neighbourhoods in the developing lexicon. *Journal of Child Language*, 30, 441–469.

Dunn, L. M., & Dunn, L. M. (2005). *Peabody Picture Vocabulary Test-III-NL*. Dutch translation by L. Schlichting. Amsterdam: Hartcourt Assessment.

Eriksson, M., Marschik, P. B., Tulviste, T., Almgren, M., Pérez Pereira, M., Wehberg, S., Marjanovič-Umek, L., Gayraud, F., Kovacevic, M., & Gallego, C. (2012). Differences between girls and boys in emerging language skills: evidence from 10 language communities. *British Journal of Developmental Psychology*, 30 (Pt 2), 326–343. doi: 10.1111/j.2044-835X.2011.02042.x.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59, 1–185. doi: https://doi.org/10.1111/j.1540-5834.1994.tb00169.x.

Fenson, L., Marchman, V. A., Thal, D., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories: User's guide and technical manual* (2nd ed.). Baltimore, MD: Brookes Publishing.

Ferrer-i-Cancho, R. (2005). The variation of Zipf's law in human language. *The European Physical Journal*, 44, 249–257.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. *Language*, 2, 301–334.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity and early word learning. In M. Bowerman & S. Levinson (eds.), *Language acquisition and conceptual development* (pp. 215–256). Cambridge, UK: Cambridge University Press.

Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language*, 35, 515–531. doi: 10.1017/S0305000907008641.

Hansen, P. (2017). What makes a word easy to acquire? The effects of word class, frequency, imageability and phonological neighborhood density on lexical development. *First Language*, 37, 205–225. doi: https://doi.org/10.1177/0142723716679956.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore:Paul H. Brookes.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, 74, 1368–1378. https://doi.org/10.1111/1467-8624.00612.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: relation to language input and gender. *Developmental Psychology*, 27, 236–248. doi: https://doi.org/10.1037/0012-1649.27.2.236.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L.V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, 61, 343–365.

Jaeger, B. (2017). R2glmm: computes R squared for mixed (multilevel) models. R package version 0.1, 2.

Jiménez, E., Haebig, E., & Hills, T. T. (2020). Identifying areas of overlap and distinction in early lexical profiles of children with autism spectrum disorder, late talkers, and typical talkers. *Journal of Autism and Developmental Disorders*. doi: https://doi.org/10.1007/s10803-020-04772-1

Johnson, P. C. (2014). Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. *Methods in Ecology and Evolution*, 5, 944–946. doi: 10.1111/2041-210X.12225.

Lewis, M., Cristiano, V., Lake, B. M., Kwan, T., & Frank, M. C. (2020). The role of developmental change and linguistic experience in the mutual exclusivity effect. *Cognition*, 198, 104191. doi: https://doi.org/10.1016/j.cognition.2020.104191.

Ma, W., Golinkoff, R. M., Hirsh-Pasek, K., McDonough, C., & Tardif, T. (2009). Imageability predicts the age of acquisition of verbs in Chinese children. *Journal of Child Language*, 36, 405–423. doi: https://doi.org/10.1017/S0305000908009008.

MacRoy-Higgins, M., Shafer, V. L., & Fahey, K. J. (2016). Vocabulary of toddlers who are late talkers. *Journal of Early Intervention*, 38, 118–129. doi: https://doi.org/10.1177/1053815116637620

McDonough, C., Song, L., Hirsh-Pasek, K., Golinkoff, R. M., & Lannon, R. (2012). An image is worth a thousand words: Why nouns tend to dominate verbs in early word learning. *Developmental Science*, 14, 181–189. doi: 10.1111/j.1467-7687.2010.00968.x.

McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science*, 317 (5838), 631. doi: 10.1126/science.1144073.

Mitchell, C., & McMurray, B. (2009). On leveraged learning in lexical acquisition and its relationship to acceleration. *Cognitive Science*, 33, 1503–1523. doi: 10.1111/j.1551-6709.2009.01071.x.

Moyle, M. J., Weismer, S. E., Evans, J. L., & Lindstrom, M. J. (2007). Longitudinal relationships between lexical and grammatical development in typical and late-talking children. *Journal of Speech, Language, and Hearing Research*, 50, 508–528. doi: 10.1044/1092-4388(2007/035).

Mulder, H., Verhagen, J., van der Ven, S. H. G., Slot, P. L., & Leseman, P. P. M. (2017). Early executive function at age two predicts emergent mathematics and literacy at age five. *Frontiers in Psychology*, 12,8: 1706. doi: 10.3389/fpsyg.2017.01706.

Pasveer, D. (2004). Test_neighbors.pl [Computer software]. Unpublished software, University of Utrecht, Utrecht, The Netherlands.

Pisoni, D. B., Nusbaum, H. C., Luce, P. A., & Slowiaczek, L. M. (1985). Speech perception, word recognition and the structure of the lexicon. *Speech Communication*, 4, 75–95. doi: 10.1016/0167-6393(85)90037-8.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reilly, J., & Kean, J. (2007). Formal distinctiveness of high- and low-imageability nouns: Analyses and theoretical implications. *Cognitive Science*, 31, 157–168.

Richardson, J. T. E. (1976). Concreteness and imageability. *Bulletin of the Psychonomic Society*, 7, 429–431. doi: 10.3758/BF03337237.

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*. 83, 1762–1774. doi: 10.1111/j.1467-8624.2012.01805.x.

Rowe, M. L., & Goldin-Meadow, S. (2009). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science*, 323, 951–953. doi: 10.1126/science.1167025.

Simonsen, H. G., Lind, M., Hansen, P., Holm, E., & Mevik, B-H. (2013). Imageability of Norwegian nouns, verbs and adjectives in a cross-linguistic perspective. C*linical Linguistics & Phonetics*, 27, 435–446.

Stokes, S. F. (2010). Neighborhood density and word frequency predict vocabulary size in toddlers. *Journal of Speech, Language, and Hearing Research*, 53, 670–683. doi: 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, 634–657. doi: https://doi.org/10.1017/S030500091300010X

Storkel, H. L. (2004). Do children acquire dense neighborhoods? An investigation of similarity neighborhoods in lexical acquisition. *Applied Psycholinguistics*, 25, 201–221. doi: https://doi.org/10.1017/S0142716404001109.

Storkel, H. L. (2009). Developmental differences in the effects of phonological, lexical and semantic variables on word learning by infants. *Journal of Child Language*, 36, 291-321. doi: 10.1017/S030500090800891X.

Storkel, H. L., & Maekawa, J. (2005). A comparison of homonym and novel word learning: The role of phonotactic probability and word frequency. *Journal of Child Language*, 32, 827–853. doi: https://doi.org/10.1017/S0305000905007099.

Storkel, H. L., & Lee, S. Y. (2011). The independent effects of phonotactic probability and neighborhood density on lexical acquisition by preschool children. *Language and Cognitive Processes*, 26, 191–211. doi: 10.1080/01690961003787609.

Swingley, D., & Humphrey, C. (2018). Quantitative linguistic predictors of infants' learning of specific English words. *Child Development*, 89, 1247–1267. doi: 10.1111/cdev.12731

van Kampen, J. (2009). The non-biological evolution of grammar: Wh-question formation in Germanic. *Biolinguistics*, 3, 154–185. doi: https://doi.org/10.1075/avt.33.05kam

van Loon-Vervoorn, W. A. (1985). *Voorstelbaarheidswaarden van Nederlandse woorden* [Imageability ratings of Dutch words]. Lisse: Swets & Zeitlinger.

Verhagen, J., Boom, J., Mulder, H., de Bree, E. H., & Leseman, P. P. M. (2019). Reciprocal relationships between nonword repetition and vocabulary during the preschool years. *Developmental Psychology*, 55, 1125–1137. doi: 10.1037/dev0000702.

Vihman, M. M., & Croft, W. (2007). Phonological development: Toward a 'radical' templatic phonology. *Linguistics*, *45*, 683–725.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, *40*, 374–408.

Wijnen, F., & Elbers, L. (1993). *Effort, production skill, and language learning. Phonological development*. Timonium, MD: York.

Zink, I., & Lejaegere, M. (2002). *N-cdi woordenlijst* [n-cdi word list]. Leusden: Acco.

Zink, I., & Lejaegere, M. (2003). *N-CDI's: Korte vormen, Aanpassing en hernormering van de MacArthur Short Form Vocabulary Checklist van Fenson et al*. [N-CDI's: Short forms, Adaptation and Re-norming of the MacArthur Short Form Vocabulary Checklist of Fenson et al.] Leuven/Leusden: Acco.

Zipf, G. K. (1936). The *psycho-biology of language: An introduction to dynamic philology*. Boston: Houghton Mifflin Company.

Appendix A: Overview of all items analyzed

egel ('hedgehog')
ezel ('donkey')
haan ('rooster')
hond ('dog')
koe ('cow')
leeuw ('lion')
olifant ('elephant')
papegaai ('parrot')
pinguin ('pinguin')
slang ('snake')
vlinder ('butterfly')
zebra ('zebra')
auto ('car')
fiets ('bicycle')
ballon ('balloon')
kleurpotloden ('crayons')
boter ('butter')
cake ('cake')
frietjes ('fries')
tafel ('table')
wasmachine ('laundry machine')
ladder ('ladder')
stok ('stick')
wolk ('cloud')
bakker ('baker')
buiten ('outside')
mevrouw ('madam')
opa ('grandpa')
kaas ('cheese')
rijst ('rice')
spaghetti ('spaghetti')
vlees ('meat')
muts ('hat')
rits ('zipper')
t-shirt ('t-shirt')
borst ('chest')
hand ('hand')
lippen ('lips')

neus ('nose')
bord ('plate')
pot ('pot')
rommel ('mess')
tandenborstel ('toothbrush')
washandje ('loofah')
zakdoek ('hankey')
slaapkamer ('bedroom')
zwart ('black')
botsen ('to bump')
kietelen ('to tickle')
knuffelen ('to cuddle')
likken ('to lick')
luisteren ('to listen')
maken ('to make')
passen ('to fit')
ronddraaien ('to spin around')
droog ('dry')
koud ('cold')
moeilijk ('difficult')
nieuw ('new')
slecht ('bad')
stout ('naughty')
verstoppen ('to hide')
zitten ('to sit')
beneden ('beneath/downstairs')
ander ('other')

Appendix B: Full Results of the Generalized Linear Mixed-Effects Models

Table B1 Results of a Generalized Linear Mixed-Effects Model Testing for Effects of Subject-Level Factors (Age, Gender, Parental Education) and Word-Level Factors (Frequency, Concreteness, Word length, Consonant Clusters, Neighborhood Density) on Acquisition

	Estimate	*SE*	*z*	*p*
(Intercept)	2.638	0.178	14.824	< .001
Age	0.226	0.019	11.711	**< .001**
Gender	0.643	0.134	4.801	**< .001**
Parental education	0.162	0.101	1.612	.107
Frequency	1.978	0.251	7.885	**< .001**
Concreteness	0.893	0.228	3.916	**< .001**
Word length	0.011	0.113	0.095	.924
Consonant clusters	−0.013	0.299	−0.045	.964
Neighborhood density	0.189	0.515	0.367	.713
Word class	0.353	0.460	0.767	.443

Note. Model: Acquired ∼ (1|Participant) + (1|Item) + Age + Gender + Parental education + Frequency + Concreteness + Word length + Consonant clusters + Neighborhood density + Word class. Number of observations: 65914; children: 1030; items: 64.

Table B2 Results of a Generalized Linear Mixed-Effects Model Testing for Effects of Subject- and Word-Level Factors, and Interactions with Age and Word Class, on Acquisition

	Estimate	SE	z	p
(Intercept)	2.709	0.222	12.218	< .001
Gender	0.643	0.134	4.817	**< .001**
Parental education	0.158	0.100	1.577	**.115**
Consonant clusters	−0.099	0.271	−0.366	.714
Age	0.209	0.020	10.479	**< .001**
Frequency	1.939	0.240	8.074	**< .001**
Concreteness	0.712	0.232	3.069	**.002**
Word length	0.063	0.102	0.619	.536
Neighbourhood density	0.187	0.460	0.405	.685
Word class	0.222	0.415	0.534	.593
Age*Frequency	−0.044	0.011	−3.862	**< .001**
Age*Concreteness	−0.033	0.007	−4.842	**< .001**
Word Class*Frequency	−1.035	0.439	−2.355	**.019**
Word Class*Concreteness	0.126	0.462	0.274	.784

Note. Model: Acquired ~ (1|Participant) + (1|Item) + Age + Gender + Parental education + Word length + Consonant clusters + Neighborhood density + (Age*Frequency) + (Age*Concreteness) + (Word class*Frequency) + (Word class*Concreteness). Number of observations: 65914; children: 1030; items: 64.

Table B3 Results of a Generalized Linear Mixed-Effects Model Testing for Effects of Subject-Level Factors, Word-Level Factors, and Interactions with Age and Vocabulary Knowledge (PPVT) on Acquisition

	Estimate	SE	Z	p
(Intercept)	2.753	0.212	12.957	**< .001**
Gender	0.409	0.133	3.066	**.002**
Parental education	−0.053	0.101	−0.524	.600
Consonant clusters	−0.150	0.255	−0.588	.556
Vocabulary	0.636	0.070	9.075	**< .001**
Frequency	1.920	0.233	8.238	**< .001**
Concreteness	0.728	0.221	3.290	**.001**
Age	0.202	0.020	10.180	**< .001**
Word length	0.089	0.096	0.930	.353
Neighbourhood density	0.303	0.439	0.689	.491
Word class	0.283	0.390	0.725	.468
Vocabulary*Frequency	−0.018	0.035	−0.504	.614
Vocabulary*Concreteness	−0.016	0.021	−0.746	.455

(*Continued*)

Table B3 (*Continued.*)

	Estimate	*SE*	*Z*	*p*
Age*Frequency	−0.050	0.012	−4.251	**< .001**
Age*Concreteness	−0.036	0.007	−4.977	**<.001**
Word Class*Frequency	−0.916	0.420	−2.180	**.029**
Word Class*Concreteness	0.177	0.436	0.407	.684

Note. Model: Acquired ~ (1|Participant) + (1|Item) + Age + Gender + Parental education + Word length + Vocabulary + Consonant clusters + Neighborhood density + (Age*Frequency) + (Age*Concreteness) + (Vocabulary*Frequency) + (Vocabulary*Concreteness) + (Word class*Frequency). Number of observations: 61307; children: 959; items: 64.