How many words do children know? A corpus-based estimation of children's total vocabulary size

Jutta Segbers and Sascha Schroeder

Max Planck Institute for Human Development, Germany

Abstract

In this article we present a new method for estimating children's total vocabulary size based on a language corpus in German. We drew a virtual sample of different lexicon sizes from a corpus and let the virtual sample "take" a vocabulary test by comparing whether the items were included in the virtual lexicons or not. This enabled us to identify the relation between test performance and total lexicon size. We then applied this relation to the test results of a real sample of children (grades 1–8, aged 6 to 14) and young adults (aged 18 to 25) and estimated their total vocabulary sizes. Average absolute vocabulary sizes ranged from 5900 lemmas in first grade to 73,000 for adults, with significant increases between adjacent grade levels except from first to second grade. Our analyses also allowed us to observe parts of speech and morphological development. Results thus shed light on the course of vocabulary development during primary school.

Keywords

Language corpus, language development, lexicon size estimation, vocabulary development, vocabulary test

Introduction

How many words does a person know? This question has interested many researchers within the last decades and led to very different approaches to solve it (e.g. Seashore & Eckerson, 1940; Goulden, Nation, & Read, 1990). However, it is difficult to answer. While the measurement of vocabulary size in young children is relatively easy since they do not know a lot of words, determining lexicon size in older children or even adults is rather challenging. Because of the amount of words they know, it is simply impossible to

Corresponding author:

Jutta Segbers, Max Planck Institute for Human Development, MPRG Reading Education and Development (REaD), Lentzeallee 94, 14195 Berlin, Germany.
Email: segbers@mpib-berlin.mpg.de

assess them all directly. As a consequence, the estimation of vocabulary size is often based on dictionaries or frequency lists: a subset of words is tested and the results are projected to the total number of words in the dictionary or list (Nation, 1993a). Other authors estimate lexicon size by analyzing (written) language production (Pregel & Rickheit, 1987). However, because of the variation in methods, estimation results differ and often only address adults' vocabulary size (e.g. D'Anna, Zechmeister, & Hall, 1991). Owing to these methodological difficulties, no reliable estimates for children's vocabulary size in primary school are available. Yet, they are necessary to describe language acquisition processes and growth rates and thus to enrich theories of vocabulary development. In addition, they enable researchers and educators to investigate different causes for vocabulary deficits in children.

In this paper, we reuse and expand existing methods to estimate children's vocabulary size on the basis of a written corpus for children (*childLex*, Schroeder, Würzner, Heister, Geyken, & Kliegl, 2015) using a corpus-based sampling approach. First, we point out the importance of vocabulary in (written) language development and describe its development and assessment. Then we present previous studies and methods to estimate total vocabulary size. Finally, we introduce our approach, its methods and results, and discuss it according to previous findings and theories on vocabulary development.

Vocabulary: Implications, development, and assessment

Vocabulary is a crucial component of language competence and language use (Nation, 1993b). It has been shown to be related to other language domains such as grammar and phonology during language development (Bates & Goodman, 1999; Gathercole & Baddeley, 1989) and is strongly connected to auditory and reading comprehension (Tannenbaum, Torgesen, & Wagner, 2006; Ouellette, 2006). Early vocabulary predicts later reading ability and school success (Muter, Hulme, Snowling, & Stevenson, 2004; Grimm & Doil, 2005), and vocabulary and reading performance stay connected throughout the lifespan (Braze, Tabour, Schankweiler, & Mencl, 2007; Landi, 2010). Thus, measuring vocabulary is common in diagnosing language impairment (Hoff, 2014). It is also necessary for further specifying the relation of vocabulary to other cognitive and language-related abilities and for conducting and planning training and intervention programs (Nation, 2012).

In describing early vocabulary development one often differentiates between receptive and productive vocabulary since language comprehension develops prior to production. Fenson, Dale, Reznick, Thal, & Pethik (1994) found that children at the age of 8 to 10 months began to understand first words. At 16 months, they comprehended more than 150 words. At the age of 12 months, children start producing their first words and are able to speak about 50 words on average at 18 months. Following that, the growth rate increases and at the age of 24 months they use about 200 words (Hoff, 2014). Vocabulary development progresses but estimates for total vocabulary size and growth for older children or young adults can rarely be found and if so, they vary substantially owing to methodological differences (see following section). However, it is commonly assumed

that receptive vocabulary exceeds productive vocabulary throughout the lifespan (Clark, Hutcheson, & van Buren, 1974). For English adults, total lexicon size is currently estimated to comprise about 50,000 words (Aitchison, 2012). Yet, it is unclear how children's vocabulary actually develops to finally reach this "goal."

As Hoff (2014) points out, early vocabulary of young children contains mostly nouns (45%). One reason for this is that they represent actual things in the children's environment; that is, they are perceptible for the child and thus their meaning is more transparent than for verbs for example. Analyzing the development of parts of speech distributions with growing vocabulary size is challenging for the same reasons lexicon size estimation itself is complicated. In their study with German school children, Pregel and Rickheit (1987) found that children's vocabularies contain about 55% nouns, about 35% verbs and 10% adjectives, based on language production of 6- to 10-year old children. They compared these numbers to Ruoff's (1981) results for adults, who estimated about 60% of nouns, 30% of verbs and 10% of adjectives. The reason for the increase in the proportion of nouns during vocabulary development is the fact that nouns are particularly likely to show effects of semantic differentiation (Clark, 1993). However, very little is known about how the prevalence of different parts of speech develops within primary school.

Another important question in vocabulary research is the nature of the relationship between lexical and morphological development. For example, Anglin (1993) found that, in English, vocabulary development is mainly driven by derivational processes and that bimorphemic words are most frequent in fifth graders' mental lexicons. By contrast, mono- and multimorphemic words are less frequent. It is unclear, however, whether these findings generalize to morphologically rich languages such as German.

In general, frequency of occurrence in a language determines which words are learned first. The more a child is exposed to a certain word the more likely he or she will be able to store it in his or her mental lexicon (Goodman, Dale, & Li, 2008; Naigles & Hoff-Ginsberg, 1998). While early language development is mostly driven by spoken language input, reading becomes more and more important in learning new words. As unknown words are more likely to appear in books (Hayes & Ahrens, 1988), the roles of print exposure and leisure time reading increase with age (Cunningham & Stanovich, 1991).

For very young children, vocabulary size is commonly tested by asking their parents to report which words out of a list their children understand (receptive vocabulary) and produce (productive vocabulary) (e.g. CDI, Fenson et al., 1993). Preschool children's vocabulary is mostly tested via picture naming (e.g. EVT-2, Williams, 2007) or picture choice after an auditory stimulus (e.g. PPVT-4, Dunn & Dunn, 2007). For school children as well as for adults, multiple–choice methods (e.g. finding a synonym out of a set of candidates) are often used regarding both their L1 and L2 vocabulary (e.g. Nation & Beglar, 2007). Another procedure introduced by Anderson and Freebody (1983) is the yes/no method, where test takers have to identify all words they know out of a list. To prevent guessing, there are pseudowords included. The authors found high correlations with actual knowledge of word meanings measured by definitions. Besides the

advantages of multiple-choice methods, the yes/no tests afford less cognitive engagement and many items can be administered within a short period of time. Since the first introduction of the method, several studies have applied it in L1- as well as in L2-language testing (Mochida & Harrington, 2006; Lemhöfer & Broersma, 2012). We introduced a German version for primary school children, which was used in this study (Segbers & Schroeder, forthcoming).

Previous studies on total lexicon size: Results and methodological issues

Many existing vocabulary tests focus on the measurement of relative vocabulary size compared to a norming sample (e.g. PPVT-4, Dunn & Dunn, 2007). Unfortunately, they provide no information on the total number of known words although these are relevant to describe vocabulary development on an average and individual level and to relate it to other developmental processes. However, various authors have tried to estimate people's total vocabulary size in different ways. Lorge and Chall (1963) distinguish between methods based on usage and sampling-based methods. In usage-based methods, spoken or written language production of the group of interest is analyzed and the number of different words is counted. With this method, Pregel and Rickheit (1987) estimated the vocabulary size of German school children aged from 6 to 10 years as consisting of up to 6900 words. However, they do not differentiate between age groups as the focus of their study was to obtain frequency norms. According to Seashore and Eckerson (1940), Marah (1872) estimated adults' vocabulary with this method to comprise from 3000 to 10,000 words. Since this approach is costly and does not provide estimates for the vocabulary size of individuals, many researchers have focused on sampling procedures. Here, a dictionary or a frequency list represents all possible known words in a language. A representative sample of words is then drawn from the dictionary or list and administered within a vocabulary test. The results are finally projected to the whole dictionary or list. One of the first attempts to estimate vocabulary size with this procedure is the study by Seashore and Eckerson (1940). They calculated a mean total vocabulary size of about 155,000 words for undergraduate college students. With the same method, Smith (1941) tested children's vocabulary size and estimated about 21,000 words for first-grade children, 38,000 words for third-grade children and 43,000 for fifth-grade children. Anglin (1993) determined a lexicon size of about 10,000 in first, 20,000 in third, and 40,000 in fifth grade and also calculated an average growth rate of 20 words per day .

The differences between the reported results are caused by some important methodological issues as described by Nation (1993a). First, the size of the dictionary or list used is crucial. According to Lorge and Chall (1963), a larger dictionary provides a better basis for vocabulary estimation since it is more likely to contain all possible words a certain person might know. Nation (1993a) points out that the dictionary has to include more words than the average test taker is believed to know to ensure that vocabulary size is not underestimated. A second and very important methodological issue is the definition of a word within the dictionary and thus within the vocabulary. Therefore, researchers have to decide whether they count derivations (e.g. *drink* vs. *drinkable*) and

inflections (e.g. *walk* vs. *walked*) as well as compounds (e.g. *main station*) as one or multiple entries. This decision reflects the assumption about representations of these morphological complex words within the mental lexicon and influences the conclusions that can be drawn from the results as well as the comparisons with other studies. Third, Nation (1993a) emphasizes the size and the compilation of the sample of items to be tested. He stresses that a larger sample leads to a smaller confidence interval for testing and thus to more accurate results. He therefore suggests that using a simple test design where a lot of items can be answered within a short amount of time without a lot of cognitive engagement. Furthermore, Nation highlights the importance of word frequency among the test items since high-frequency words are more likely to be known. He suggests ordering words by frequency classes of the same size and then taking the same number of words from every frequency level so that neither high- nor low-frequency words are overrepresented within the sample. Finally, Nation points out the necessity for authors to report clearly all the decisions described above, so that other researchers can evaluate and replicate the findings. In a later review, Nation (2012) advises, owing to technological progress and the emergence of language corpora, to prefer a frequency-based sampling over the dictionary-based method. He therefore suggests building up a corpus that contains a representative sample of the words of the language of interest. For German read by children, such a corpus was introduced by Schroeder and colleagues (2015) and will be described below.

The childLex corpus and the German language

The *childLex* corpus (Version 0.16.03; Schroeder et al., 2015) is a written language corpus for German read by children and contains linguistic data for words from 500 children's books. It comprises about 10,000,000 tokens, 180,000 types and 117,000 lemmas. The corpus was intended to include books that are frequently read by children aged 6 to 12 years in school and in their leisure time. Both teacher and children questionnaires and library lending statistics were considered as part of the book selection process. Thus, we assume that it is representative for the written language exposure of German school children and that the relative frequencies of the corpus can be used to approximate the order in which words are learned (Naigles & Hoff-Ginsberg, 1998). It therefore meets the criteria for the basis of vocabulary size estimation described by Nation (1993a, 2012).

As Nation (1993a, 2012) pointed out, the definition of the unit of analysis is crucial for vocabulary size estimation. In contrast to English, German is a morphologically rich language (Fleischer, Barz, & Schröder, 2012). Concerning its inflection, for example, a verb such as *lachen* ("to laugh") can appear in 13 different forms depending on person and tense (*ich lache, du lachst,* etc.). In comparison, in English there do exist four different forms of the word (*laugh, laughs, laughed, laughing*). Nouns and adjectives are also inflected according to number and case in German. Furthermore, German is a very productive language. Especially compounding is very common and, in contrast to English, compounds are mostly written without spacing (e.g. *Bahnhof* means *train station*). Also, derivation is very frequent in German, e.g. the prefix "un-" can be

combined with adjectives to form an antonym (e.g. *glücklich – unglücklich*, *happy – unhappy*). While inflection in German is supposed to happen post-lexically, it is unclear whether compositions and derivations are stored as whole units within the mental lexicon or combined after retrieval of the single constituents (Fleischer et al., 2012). We therefore decided to use the lemma as the base unit of our analysis. In the following, a lemma is defined as the abstracted base form of a word. Thus, all inflectional forms of a word are represented by the same lemma whereas compounds and derivations are counted as different lemmas. D'Anna and colleagues (1991) argue that a lemma represents a base word in a language and thus is the best count for different words known. Thus, the 117,000 lemmas of the childLex corpus served as the basis for our vocabulary estimation method. Due to the fact that a lot of words in a language do only occur very infrequently (e.g.in childLex, 48.30% of lemmas occur only once within the corpus), a frequency-level classification and sampling scheme as suggested by Nation is not feasible: A lot of very infrequent words would have to be tested to project the results to the whole corpus as it was done in previous studies. We therefore decided to apply a sampling-based method which allowed us to draw item and person characteristics from the corpus.

Our approach: A corpus-based estimation of vocabulary size

In the present study, we estimate the vocabulary size of school children at different ages and of young adults. We created a vocabulary test based on the yes/no method introduced by Anderson and Freebody (1983). To determine total vocabulary size, we then reused and expanded the dictionary method described by Nation (2012) using the childLex corpus as our basis. Based on the assumption that the relative frequencies in childLex are representative for children's written language exposure, we drew virtual lexicons of different sizes from the corpus and let them "take" a vocabulary test. To this end, we repeatedly sampled different lexicon sizes from the corpus and checked whether or not the test items were included in the lexicon. Thus, given a specific lexicon size, we know the probability that a particular test item can be solved. This allowed us to identify the relation between test results and lexicon size. We then let a real sample of German school children and young adults take our vocabulary test and used the results from the virtual dataset for the estimation of participants' total vocabulary size. Our method also enabled us to compute vocabulary growth rates and to estimate the development of parts of speech proportions and morphological categories within the mental lexicon.

Method

Sample

A sample of 495 children took part in the study. Twenty-four children (4.84%) did not complete the vocabulary test, resulting in a final sample of 471 children (249 female, 217 male, five not reported). Children's data were collected in primary (grades 1 to 6) and

Table 1. Sample characteristics.

Grade	N	M age	Gender		
			Male	Female	NA
1	37	6.6 (0.5)	13	23	1
2	49	7.3 (0.7)	24	25	0
3	75	8.0 (0.6)	38	35	2
4	107	9.0 (0.6)	65	42	0
5	62	10.0 (0.6)	22	40	0
6	68	11.2 (0.6)	35	33	0
8	73	12.8 (0.5)	21	51	1
Adults	30	22.4 (2.1)	14	16	0

Note: SD provided in parentheses.

secondary (grade 8) schools in Berlin. Thirty young adults (16 female, 14 male) were tested at the Max Planck Institute for Human Development Berlin. The study was approved by the ethics committee of the Max Planck Institute for Human Development and by the school administration of Berlin. Participation was voluntary and based on parental consent if necessary. Children received candy for their participation and adults were reimbursed with 12 Euros. The number of participants in each grade as well as mean age and gender distribution are provided in Table 1.

Vocabulary Test

The self-developed Vocabulary Test was based on the yes/no method introduced by Anderson and Freebody (1983; Segbers & Schroeder, forthcoming). In this test, participants were presented with a list of 100 words and had to identify all words they knew. To prevent guessing, the list also contained 24 pseudowords. We created five test versions for different age groups (first/second grade, third/fourth grade, fifth/sixth grade, eighth grade, adults). Because item difficulty mainly depends on word frequency, we decreased mean log lemma frequency systematically in order to ensure an optimal level of test difficulty in each version (see Table 2). To link the different test versions, a subset of 20 words was used in all age groups and subsequent test versions each shared 10 overlapping link items. Owing to a technical error, there were 11 shared items between Version 5/6 and Version 8. Thus, the total number of items was 379. The number of items in each test version is provided in Table 2. Pseudowords were created by exchanging the vowels of a different list of real words (e.g. *schwach* to *schwich*) or by combining two existing morphemes (e.g. *Führtum*) and were identical in all test versions. For each version, two randomized pseudoparallel forms A and B with different word orders were created. Participants were instructed to identify all known words. They were told explicitly that the list also comprised pseudowords and that thus guessing could easily be detected. Depending on participant's age, the test took between 5 and 15 minutes.

Table 2. Frequency distribution and number of items in different test versions.

Test version	Log Lemma Frequency		N items*		
	M (SD)	Range	Unique	Overlap	Link
1./2. Grade	1.5 (0.4)	2.7 – 1.0	70	10	20
3./4. Grade	0.6 (0.1)	0.9– 0.4	60	10	20
5./6. Grade	0.0 (0.1)	0.4 – −0.2	59	11	20
8. Grade	−0.3 (0.2)	0 – −0.6	59	10	20
Adults	−0.7 (0.2)	−0.1 – −1.0	70	–	20

*"Unique" indicates the number of items that appear only in the test version of this age group. "Overlap" indicates the number of items that are shared between the test version of this age group and the next higher age group. "Link" indicates the number of items shared between all test versions.

Analysis

All analyses were performed with the Software R (R Core Team, 2015). Data analysis comprised of a sequence of four interconnected steps: First, we drew virtual samples of different lexicon sizes from the corpus. Second, we let the virtual samples "take" the vocabulary test and estimated corresponding item and person parameters using item response theory (IRT). This allowed us to determine the relationship between lexicon size and person parameters. Third, we used the virtual item parameters to estimate a person parameter for each participant in our empirical sample. As the relationship between person parameters and lexicon size is known, it is therefore possible to estimate an individual's vocabulary size and compute growth rates between age groups. In a final step, we additionally analyzed the development of different parts of speech and morphological categories. Each step is explained in detail in the Results section.

Results

Sampling of lexicon sizes

In a first step, we sampled virtual lexicons of different sizes from the childLex corpus. The sampling procedure was based on the list of all lemmas included in childLex (version 0.16.03; approx. 117,000 lemmas) and sensitive to the frequency of each lemma, i.e., high-frequency lemmas were more likely to be included in a sample than low-frequency lemmas. We varied lexicon sizes from 1000 to 115,000 lemmas. Between 1000 and 70,000 lemmas, lexicon size was increased in steps of 100. Between 70,000 and 115,000 lemmas, lexicon size was increased in steps of 5000 lemmas. This resulted in 700 different lexicon sizes which were sampled 100 times each by drawing the according number of lemmas from the corpus. All 70,000 virtual lexicons were used for further analyses.

Owing to the sampling procedure, virtual lexicons of the same size could potentially comprise completely different lemmas. However, since sampling was based on word frequency, high-frequently used lemmas were more likely to be included in several lexicons at the same time. As a consequence, small lexicon sizes shared a substantial proportion of their lemmas with each other while large lexicon sizes were more heterogeneous.

Thus, small lexicon sizes were more likely to contain high-frequency lemmas but could also comprise lemmas with low frequencies. With growing lexicon size the amount of low frequent lemmas increased.

Estimation of virtual item and person parameters

In a next step, we examined whether the 379 words of our vocabulary test were included in each virtual lexicon or not. In other words, we let our virtual lexicons "take" each of the five versions of the vocabulary test by comparing the sampled lexicons with our test items. If a test item was included in the lexicon, we considered it as known by the virtual participant with the according lexicon size. In case a test item was not included in the lexicon, we assumed that is was not known by this virtual participant. Thus, we were able to compute the probability that a particular test item can be solved as a function of the size of a lexicon.

The relationship between item solving probability and lexicon size is just a special case of the dependency between person ability and item difficulty. In order to analyze such relationships, item-response models are ideal and thus commonly used in educational testing. Model estimation was executed with the ltm package for R (Rizopoulos, 2006). In the present study, we analyzed our virtual data using the two-parameter-item-response model (2 PL-Model; Embretson & Reise, 2000; Bock & Zimowski, 1997). In the 2 PL-Model, the two item parameters of difficulty and discrimination as well as the person parameter representing the latent ability are estimated. The model fitted our data, $LL = -8307611$, and a comparison with the simpler one-parameter model ($LL = -8437834$) revealed a significant improvement of model fit for the two-parameter model, $\Delta \chi^2 (378) = 260445$, $p < .01$.

Because the fit of the model was adequate, we fitted item parameters using Conditional Maximum Likelihood (CML) and saved them for further analyses. All 70,000 virtual lexicons were included in the analysis. Item difficulty ranged from −4.85 and 3.19 ($M = -0.36$, $SD = 1.98$); its distribution is displayed in Figure 1 A. Item discrimination ranked between 0.79 and 4.18 ($M = 1.63$, $SD = 0.78$) and is presented in Figure 1 B.

In addition, person parameters for the different lexicon sizes were calculated via expected-a-priori (EAP) estimation using the PP package for R (Reif, 2014). Estimated person parameters ranged from −4.10 to 5.52 ($M = 0.03$, $SD = 1.63$).

Because lexicon size (and, therefore, latent ability) was known a priori, this enabled us to relate person parameters and virtual lexicon sizes (see Figure 2). The relationship was very strong and could nearly perfectly be captured by a cubic function ($R^2 = .99$). Since we know the relation between person parameters and lexicon size, we are able to transform person parameters to lexicon sizes (and vice versa). This procedure can also be applied to person parameters derived from empirical samples.

Estimation of empirical person parameters and lexicon sizes

In a last step, data from our empirical sample was analyzed. Since we know the relationship between item difficulty and person ability (see above), virtual item parameters were used to examine participants' actual performance on the Vocabulary Test and to
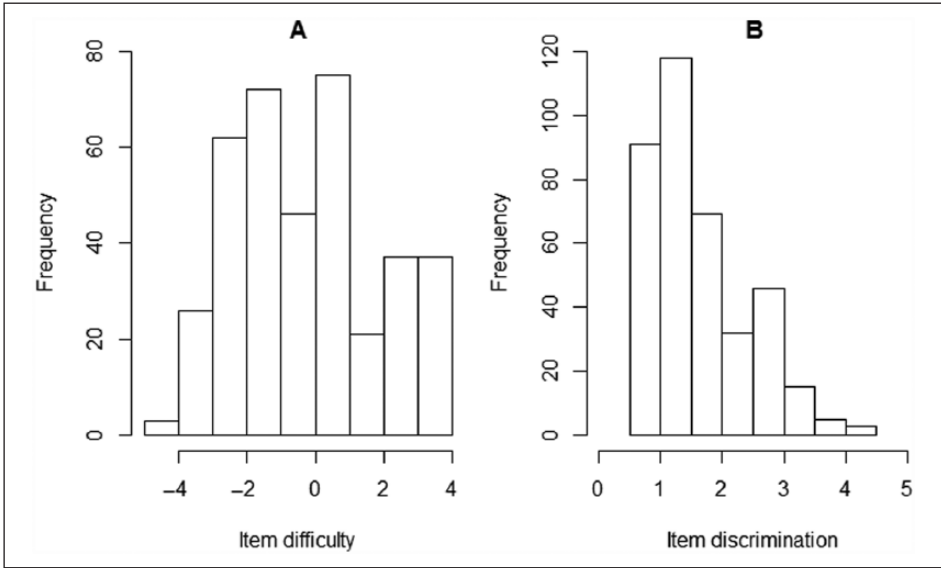
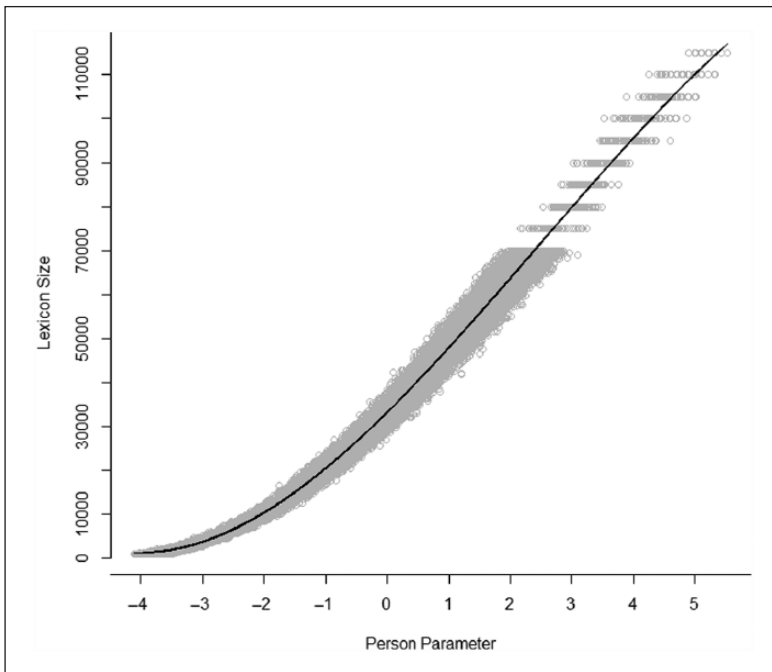**Figure 1.** Distribution of item difficulty (A) and item discrimination (B) derived from the virtual lexicons.

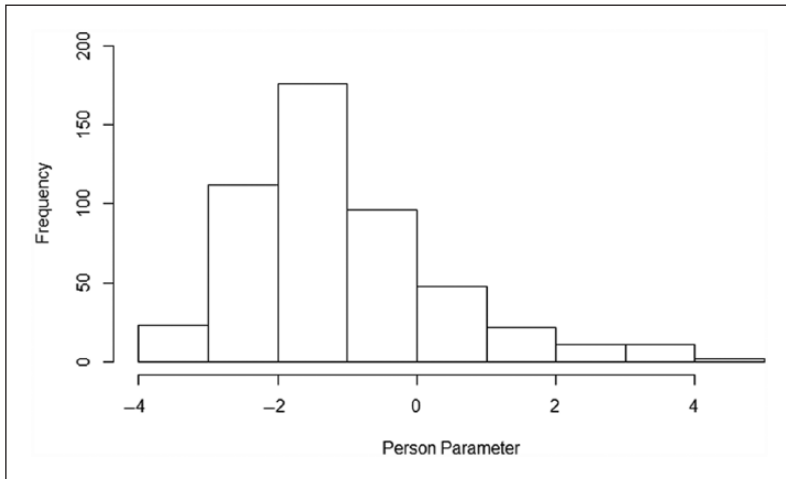**Figure 2.** Relation of lexicon size and person parameter.

Figure 3. Distribution of person parameters in the empirical sample.

Table 3. Mean lexicon sizes and standard deviations per grade.

Grade	M person parameter (SD)	M lexicon size (SD)
1	−2.66 (0.44)	5925 (2481)
2	−2.68 (0.61)	6097 (4063)
3	−1.96 (0.54)	11,182 (4647)
4	−1.56 (0.54)	14,819 (5643)
5	−1.19 (0.60)	18,812 (6796)
6	−0.61 (0.67)	25,694 (8706)
8	0.30 (0.77)	38,029 (11,107)
Adults	2.60 (1.14)	73,625 (17,593)

estimate corresponding person parameters. Again, parameter estimation was based on EAP. The distribution of person parameters is provided in Figure 3. The overall mean was $M = -1.12$ ($SD = 1.46$) and parameters ranged between −3.92 and 4.51. Mean person parameters and standard deviations per grade are provided in the first column of Table 3.

To investigate whether the item parameters derived from the virtual lexicons were appropriate to fit our real data, we also estimated person parameters using item parameters derived from the empirical sample directly. The person parameters from both analyses correlated highly ($r = .93$) indicating that using the virtual item parameters was appropriate.

Finally, we were able to transform the empirical person parameters into individual lexicon sizes via the cubic function derived above. Mean lexicon sizes and standard deviations for each grade are provided in the second column of Table 3. As expected, our
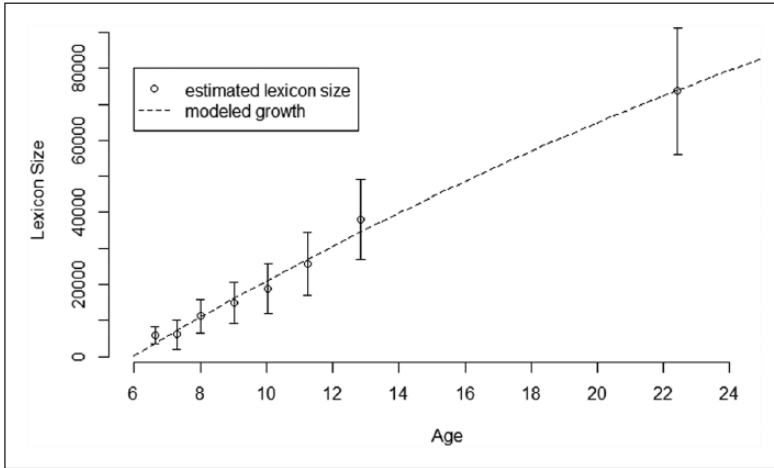
Figure 4. Development of lexicon size by age with modeled quadratic function (bars represent standard deviations).

lexicon size estimation shows a growing trend, with about 6000 lemmas in first grade and about 73,000 lemmas in young adults.

Figure 4 shows the growth of lexicon size between grades by plotting average lexicon size per grade against mean age per grade. An ANOVA revealed a significant effect of *grade* on lexicon size, $F(7, 493) = 302.8$, $p < .05$. Post-hoc analyses showed no significant difference between first and second grade. All other differences were significant (all $p < .01$). The high standard deviations, however, point to great interindividual differences within each grade. Generally, vocabulary growth could nearly perfectly be described as a quadratic function of age, $R^2 = .99$, also displayed in Figure 4:

$$Lexicon\ Size\ =\ -54.59 * Age^2\ +\ 6032.58 * Age\ -\ 33938.01 \tag{1}$$

Further analyses: Parts of speech development and vocabulary growth rates

Our estimation method also allowed us to analyze the development of different parts of speech within the vocabulary (e.g. nouns, verbs and adjectives) as childLex contains parts-of-speech tagging. The sampling of different virtual lexicon sizes from childLex enabled us to count word classes within the lexicon. The investigated categories were *nouns, verbs, adjectives, function words*, and *others* (containing, e.g., proper names). The development of different word classes with growing lexicon size is displayed in Figure 5. Clearly, the number of words is increasing in every category (Figure 5 A).
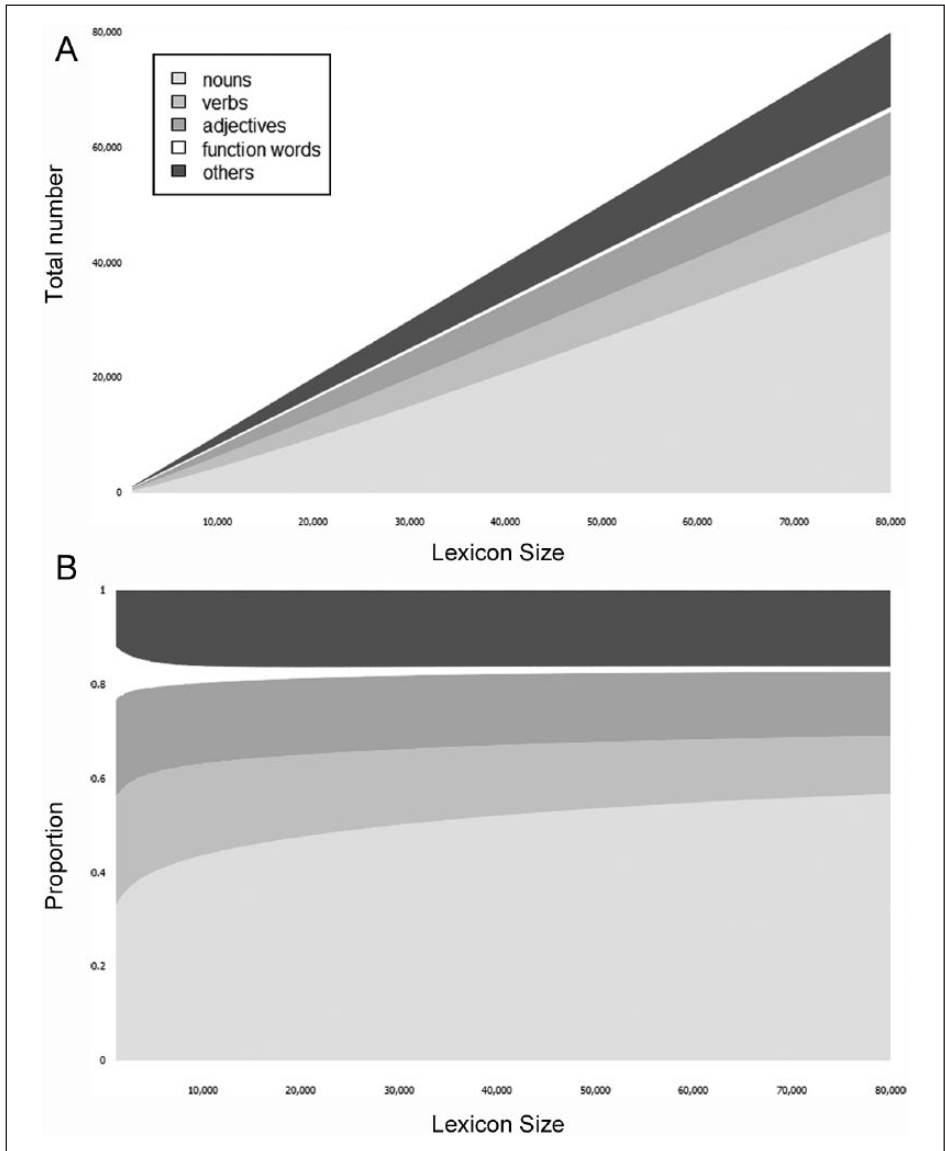
Figure 5. Development of parts of speech with lexicon size in total numbers (A) and proportions (B).

Regarding the proportions of the different parts of speech, only the noun category is growing while the proportions for the other parts of speech are slightly decreasing with growing lexicon size (Figure 5 B).

Table 4. Mean numbers of parts of speech per grade.

Grade	Nouns	Verbs	Adjectives	Function words	Others
1	1697 (1426)	1655 (296)	1286 (338)	327 (20)	960 (401)
2	1795 (2335)	1676 (486)	1309 (554)	329 (33)	988 (656)
3	4718 (2671)	2284 (555)	2002 (633)	369 (37)	1,809 (750)
4	6808 (3243)	2718 (674)	2498 (769)	399 (45)	2396 (911)
5	9103 (3906)	3195 (812)	3042 (926)	431 (55)	3041 (1097)
6	13,058 (5003)	4018 (1040)	3980 (1186)	486 (70)	4152 (1406)
8	20,147 (6383)	5492 (1327)	5661 (1514)	585 (89)	6144 (1794)
Adults	40,604 (10,111)	9746 (2103)	10,512 (2398)	871 (141)	11,892 (2841)

Note: SD provided in parentheses.

Table 5. Estimated average vocabulary growth rates (in words per day).

	Parts of speech					Total
	Nouns	Verbs	Adjectives	Function words	Others	
Grade 1 – Grade 2	0.27	0.06	0.06	0	0.08	0.47
Grade 2 – Grade 3	8.01	1.67	1.90	0.11	2.25	13.93
Grade 3 – Grade 4	5.73	1.19	1.36	0.08	1.61	9.96
Grade 4 – Grade 5	6.29	1.31	1.49	0.09	1.77	10.94
Grade 5 – Grade 6	10.83	2.25	2.57	0.15	3.04	18.85
Grade 6 – Grade 8	9.71	2.02	2.30	0.14	2.73	16.90
Grade 8 – adults	14.01	2.91	3.32	0.20	9.94	24.38

By identifying the function between lexicon size and parts of speech displayed in Figure 5 A, we were able to estimate the average numbers of parts of speech within the vocabularies of our empirical sample (Table 4). Results show that nouns dominate vocabulary at all stages of lexical development, followed by verbs, adjectives, other words (e.g. proper names), and function words.

Similar to Anglin (1993), we also calculated average vocabulary growth rates, that is the number of words learned per day. To this end, we computed the difference between grades for both total lexicon size and for the different word categories and divided it by the average age difference between grades (assuming one year per grade and 365 days per year). Results are displayed in Table 5. Whereas growth rates from first to second grade are relatively small, they range between about 10 and 20 words per day from grade 2 to grade 8. Again, most of children's vocabulary growth is driven by learning new nouns.

Similarly, we examined the development of different morphological categories within our sample. To this end, all lemmas in childLex were analyzed using the morphological tagger SMOR (Schmid, Fitschen, & Heid, 2004). Here, we concentrate on two different variables: Morphemic complexity (mono-, bi- or multimorphemic, i.e. words that consisted of three or more morphemes) and morphological category (monomorphemic, derivation, composition, derivation & composition). Since our analyses are based on lemmas, inflection was not included in the morphological categorization. Figure 6 shows the
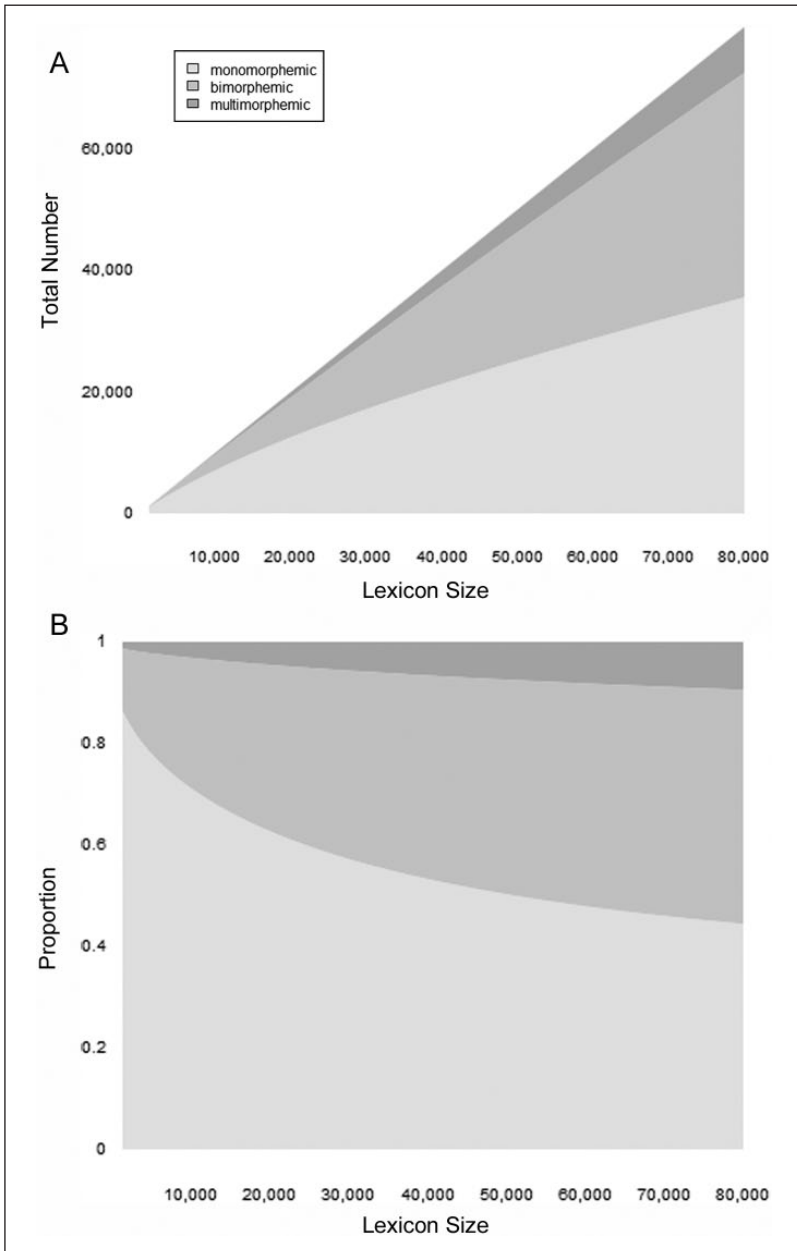
Figure 6. Development of morphological complexity with lexicon size in total numbers (A) and proportions (B).

development of words with different morphemic complexity with increasing lexicon size. Clearly, the number of lemmas in each category increases (Figure 6 A). However, the corresponding proportions (Figure 6 B) show that the percentage of monomorphemic

Table 6. Means of numbers of words by morphemic complexity per grade.

Grade	Monomorphemic	Bimorphemic	Multimorphemic
1	4865 (1334)	974 (1011)	87 (136)
2	4940 (2131)	1054 (1690)	102 (243)
3	7624 (2396)	3159 (1960)	399 (293)
4	9478 (2809)	4706 (2443)	635 (393)
5	11,475 (3330)	6439 (2995)	916 (504)
6	14,730 (3971)	9514 (4001)	1450 (742)
8	20,202 (4693)	15,274 (5340)	2552 (1084)
Adults	33,005 (6052)	33,785 (9373)	6835 (2259)

Note: SD provided in parentheses.

Table 7. Means of morphological word types per grade.

Grade	Monomorphemic	Derivation	Composition	Derivation + Composition
1	4865 (1334)	602 (410)	427 (679)	32 (57)
2	4940 (2131)	628 (665)	490 (1165)	38 (104)
3	7624 (2396)	1462 (756)	1931 (1372)	165 (126)
4	9478 (2809)	2052 (906)	3023 (1,761)	267 (170)
5	11,475 (3330)	2692 (1,082)	4275 (2,199)	388 (219)
6	14,730 (3971)	3780 (1357)	6563 (3063)	620 (325)
8	20,202 (4693)	5685 (1688)	11,037 (4259)	1104 (478)
Adults	33,005 (6052)	10,840 (2511)	26,770 (8126)	3010 (1008)

Note: SD provided in parentheses.

lemmas consistently decreases while the percentage of bi- and multimorphemic lemmas increases with growing lexicon size. Thus, most of children's lexicon growth is driven by the acquisition of morphologically complex words.

The same trend can be seen in Table 6, which provides the average number of words in each complexity category separately for each grade level. While the number of bi- and multimorphemic lemmas is relative small in first-graders, it strongly increases during lexical development. In adults, bi- and multimorphemic words constitute the majority of words in the lexicon.

To analyze further the development of morphological complex words, we classified the bi- and multimorphemic lemmas into different categories depending on whether they are formed by derivation, composition or a combination of both processes. Figure 7 shows the development of the different categories with increasing lexicon size. Although all categories generally increase during lexical development, the relative growth is most pronounced for compound words.

Table 7 shows the means of different morphological categories for each grade separately. In grade 1 and grade 2, derivation is the most prominent morphological category. After this, however, compounds become more frequent than derivations. The proportion of combinations of derivation and composition is generally relatively small.
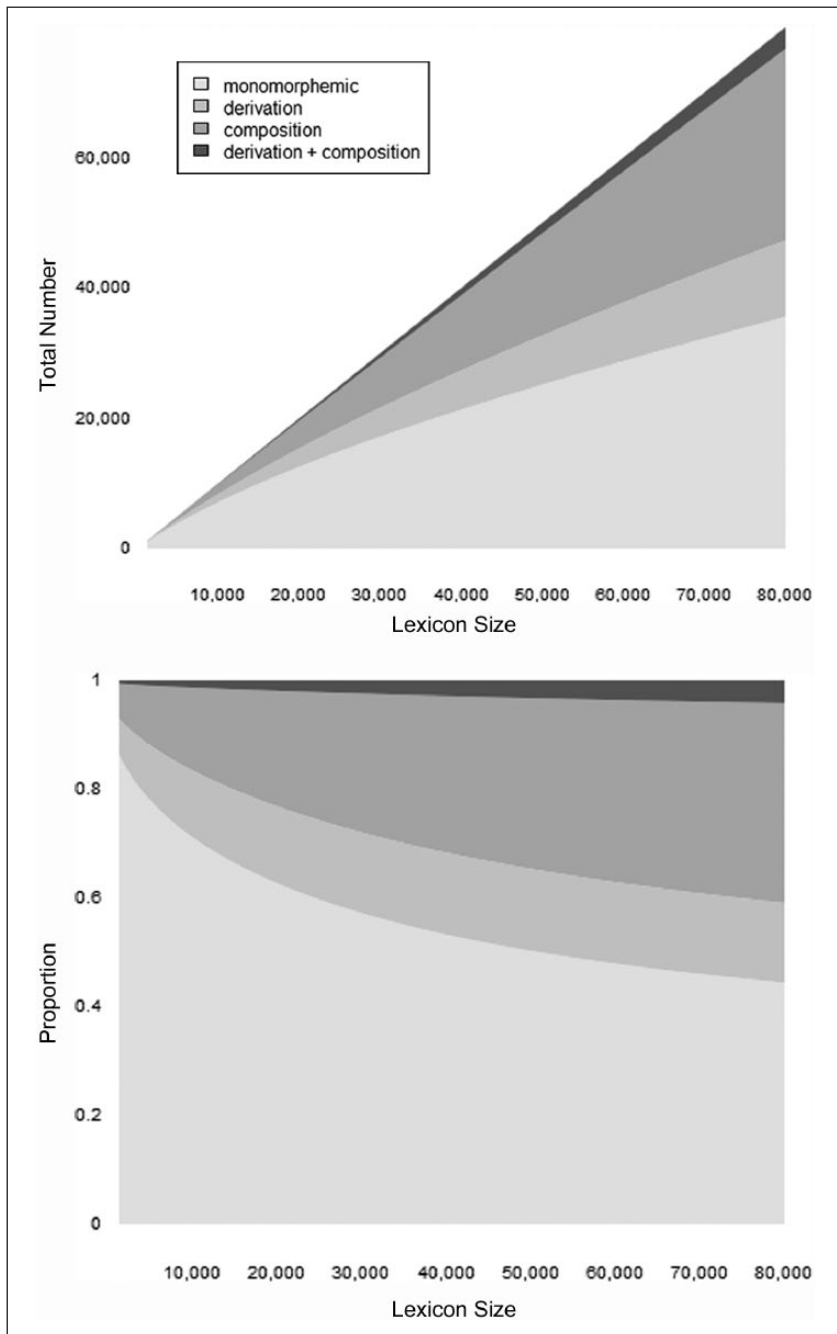
Figure 7. Development of morphological word types with lexicon size in total numbers (A) and proportions (B).

Table 8. Growth rates for the morphological categories in words per day.

	Monomorphemic	Derivation	Composition	Derivation + Composition
Grade 1 – Grade 2	0.21	0.07	0.17	0.02
Grade 2 – Grade 3	7.35	2.29	3.95	0.35
Grade 3 – Grade 4	5.08	1.61	2.99	0.28
Grade 4 – Grade 5	5.42	1.75	3.43	0.33
Grade 5 – Grade 6	8.97	2.98	6.27	0.64
Grade 6 – Grade 8	7.50	2.61	6.13	0.66
Grade 8 – adults	8.77	3.53	10.78	1.31

Average growth rates (words/day) for the different morphological categories are provided in Table 8. Results again show that vocabulary development is mostly driven by monomorphemic words in early grades, but morphologically complex words and particularly compounds are becoming increasingly important during later lexical development.

Discussion

In this study, we introduced a corpus-based method to estimate children's vocabulary size. We used a vocabulary test based on the yes/no method (Anderson & Freebody, 1983). From the childLex corpus, we drew a virtual sample of different lexicon sizes and estimated item and person parameters for our test with this sample by using the 2 PL-IRT-Model (Embretson & Reise, 2000). This enabled us to identify the function between person parameters and lexicon size. We then let a real sample of school children from grades 1 to 6, 8 and young adults take the test. Person parameters were estimated using item difficulties derived from the virtual sampling procedure. As a consequence, we were able to compute the total lexicon sizes for each participant. Based on these estimates, we analyzed average vocabulary sizes and growth rates per grade as well as the proportion of different parts of speech and morphological categories with increasing lexicon size.

Our method is an extension of the previously used dictionary methods and meets the criteria for vocabulary size estimation described by Nation (2012). It is based on a corpus which is representative for the frequency structure of German read by children. Furthermore, we used a vocabulary test which contains a lot of items but requires only little cognitive effort. The fit of the 2 PL-Model was satisfying and the correlation of person parameters derived from the virtual item parameters with person parameters derived from the real data was high. We therefore think it is a useful method for the estimation of people's total vocabulary size.

Our estimation of school children's vocabulary fills the gap between the known vocabulary size of young children and the assumptions about adult lexicon size in German. Our results show lexicon size increases from approximately 6000 lemmas in first grade to about 73,000 lemmas in young adulthood. In comparison to previous studies on total vocabulary size in English children (see, e.g., Anglin, 1993), our estimates are considerably smaller. These differences may be due to methodological disparities: Anglin's estimation, for example, included root words as well as inflected and derived

words, literal compounds and idioms. By contrast, our estimation was based only on lemmas, thus inflected words are excluded. When the number of inflected words is subtracted from total lexicon size, the estimates reported by Anglin (1993) and here are more similar (and the same holds for the study of Smith, 1941). Relatedly, the estimates provided by our method are substantially larger than the estimates that have previously been reported for German school children (e.g. Pregel & Rickheit, 1987). Again, these discrepancies might be driven by methodological differences. While Pregel and Rickheit (1987) analyzed written and spoken language production, our method is based on a receptive vocabulary measure. Since receptive vocabulary is assumed to be larger than productive vocabulary (Clark et al., 1974), our results do not contradict but complement the findings of Pregel and Rickheit by also providing estimates for children's receptive vocabulary development. With regard to adults' vocabulary size, our estimate of 73,000 lemmas exceeds the commonly assumed number of 50,000 words in English (Aitchison, 2012). This difference might be ascribed to cross-linguistic disparities: German is a morphologically rich language and compounding is particularly frequent. As a consequence, the number of different words or lemmas is likely to be higher than in English. This, in turn leads to larger vocabulary size estimates for German adults. Surprisingly, we did not observe the same language effect for children's vocabularies which were generally smaller, but not larger than in English. However, none of the previous studies has tried to estimate the development of vocabulary from child- to adulthood as it was done in the present investigation, leading to comparable estimates for all age groups as in our study. This is important, because our results for the growth rates of different parts of speech and morphological categories showed that most of children's vocabulary growth was related to the acquisition of nouns and particularly compounds. This is in line with findings on early vocabulary acquisition in German (Hoff, 2014; Pregel & Rickheit, 1987).

With regard to vocabulary growth, we found strong differences between almost all age groups indicating a remarkable development of vocabulary from primary school to adulthood. Only between first and second grade, no significant increase in vocabulary size was observed. This finding might be explained by the fact that children are still learning to read during that time. Vocabulary growth in school is mostly driven by reading activity since new words are more likely to occur in written than in spoken language (Hayes & Ahrens, 1988; Nagy, Herman, & Anderson, 1985). During first grade, children's reading ability is at a low level, thus their reading input is relatively small, leading to limited vocabulary growth. After this initial phase, however, children's vocabulary grows by several thousand lemmas a year as has been reported in other studies (e.g. Anglin, 1993; Smith, 1941). Overall, the developmental trajectory could well be described by a quadratic function which is in line with prior findings on early vocabulary development (Kauschke & Hofmeister, 2002; Huttenlocher, Haight, Byrk, Seltzer, & Lyons, 1991).

The accelerating dynamic of children's vocabulary growth is also demonstrated in our average growth rates of learned words per year. Again, we only find small growth rates in grade 1. From grade 2 to grade 8, however, growth rates are substantial and vary between 10 and 20 learned words per day. After grade 8, growth rates are even higher with approximately 25 new words per day. Thus, vocabulary growth is not completed after the end of children's compulsory school education but is likely to increase owing to further education and experiences during young adulthood. Future research on older adult's vocabularies using our estimation method could shed light on the determinants of this process.

It is important to note that the standard deviations in all age groups were very high indicating substantial variability in vocabulary size. In grade 4, for example, low-performing children (−1SD) have functional vocabulary sizes that are similar to the average vocabulary size in grade 2 and high-performing children (+1SD) have vocabulary sizes that are similar to the average vocabulary size in grade 6. This highlights the importance of investigating interindividual differences in vocabulary development. In addition, it emphasizes the necessity to be able to assess the vocabulary size of each participant individually in language assessments, and not to rely on grade-level averages.

Our findings also confirm another crucial point: Given the enormous growth rates per year, vocabulary cannot only be taught in school alone (Nation, 1993b). As Jenkins, Stein, and Wysocki (1984) suggested, other activities such as leisure time reading, practicing hobbies or watching movies are similarly important. According to Nation (1993b), teachers should encourage their students to engage in such activities and thus support indirect vocabulary learning. Regarding the growth rates of parts of speech within the lexicon, we found that vocabulary development is mostly driven by the increasing number of known nouns. This leads to further implications for vocabulary growth and vocabulary teaching: Teaching nouns, directly or indirectly, plays an important role in supporting vocabulary growth.

Our investigation of morphological development showed children's lexical development is strongly driven by the acquisition of morphologically complex words. In early grades, most newly acquired complex words are derivations but in later grades the acquisition of compounds dominates lexical development. Anglin (1993) also observed a decline in the proportion of monomorphemic words and an increasing percentage of bi- and multimorphemic words with vocabulary development in English. However, in his study this trend was mostly driven by the acquisition of new derivations but not compounds. Again, this discrepancy might partly be explained by cross-linguistic differences. Compounding is very frequent in German and most compounds can be generated and understood spontaneously (e.g.the non-existing compound "Kleinkinderbaum" [small children's tree] can easily be interpreted as a kind of tree that has been especially designed, planted, etc. for small children). This example demonstrates that orthographic and methodological differences might also contribute to the diverging findings between English and German. In contrast to German, compounds are usually written using spaces or hyphens in English. As a consequence, they are not recognized by algorithms that solely rely on white-space segmentation for tokenization (Jurish & Würzner, 2013). In sum, our findings fit to the assumption that morphological processes become more and more important during vocabulary development. This has important educational implications and shows that it is essential to call attention to morphological processes such as derivation and composition to enrich the learner's vocabulary.

In the present study, we used lemmas as the basic unit of analysis as it is particularly suited for an inflectionally rich language such as German. However, the method can be easily applied to other linguistic entities (e.g.inflected word types or stems) depending on the assumption on how words are stored in the mental lexicon. For example, if one assumes that inflected word forms constitute distinct lexical entries, the virtual sampling procedure as well as the selection of test items would simply be based on this unit of analysis. Similarly, the method can easily be extended to other corpora (e.g. for adults or other languages).

Limitations and challenges

Although our approach appears very promising there are obviously also some limitations and challenges concerning the interpretation of our results and the application of the method for further research. First, we have to point out that our method crucially depends on the corpus which is used. As Kornai (2002) emphasizes, a language does not contain a finite number of words, mostly because of productive morphological processes which are especially important for languages such as German. In addition, as any corpus is only sample, its quality depends on its representativeness of the unobserved target population. This issue is particularly important with regard to frequency counts since our sampling method is sensitive to the frequency distribution in the corpus. In this regard, Fengxiang (2010) stresses that especially the number and frequency of rare words varies substantially with corpus size and might be underestimated.

We believe that the childLex corpus generally meets the necessary requirements regarding both quantity and quality. However, it also has some obvious limitations. The corpus mostly comprises narrative but not expository texts. Since children spend much more time reading narratives as opposed to non-narratives, this is consistent with the goal of representing the words that children have most likely encountered in their leisure-time reading (Topping, 2015). However, this necessarily limits its predictive value for students' performance on expository reading assignments which are more common in school settings. Furthermore, we also used the childLex frequencies to estimate adults' vocabulary size. Although there is a strong relationship between the frequency counts in childLex and corresponding adult corpora (Schroeder, Würzner, Heister, Geyken, & Kliegl, 2015), there are also clear discrepancies, especially for low-frequency words. As a consequence, the results for the adults might be less precise than for children. In summary, the methods employed for corpus construction are crucial for the application of our vocabulary estimation method and thus have to be carefully evaluated before using the approach.

We also have to emphasize that our method leads only to an approximation of total lexicon size and that there might be other factors than word frequency that drive vocabulary acquisition which were not considered here. Our estimation does not involve semantic information, for example concerning homonymy or polysemy. Also, our results do only contain information about the number of known words, and not on the quality with which the words are represented within the mental lexicon (Perfetti & Hart, 2002).

Conclusions

In conclusion, the reported findings make an important contribution to the discussion about vocabulary development in German across the school years. The proposed method cannot only be used to estimate total vocabulary size but also allows analyzing other linguistic phenomena such as the development of different parts of speech or morphological categories within the mental lexicon. The resulting estimates of total lexicon size are useful for describing and understanding language acquisition processes in German. Most importantly, in contrast to other methods, our approach enables researchers to estimate the lexicon size for each individual separately. Thus, differential developmental trajectories and their effects on children's reading performance can be investigated on an

individual level. This is important in order to gain further insights about the relation between vocabulary and cognitive development, which might be used to improve existing training and intervention methods.

Funding

References

Aitchison, J. (2012). *Words in the mind: An introduction to the mental lexicon* (4th ed.). Malden, MA: Wiley.

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching*. Newark, DE: International Reading Association.

Anderson, R. C., & Freebody, P. (1983). Reading comprehension and the assessment and acquisition of word knowledge. *Advances in Reading/Language Research*, *2*, 231–256.

Anglin, J. M. (1993). *Vocabulary development: A morphological analysis. Monographs of the Society for Research in Child Development*, *58*. Hoboken, NJ: Wiley.

Bates, E., & Goodman, J. C. (1999). On the emergence of grammar from the lexicon. In B. MacWhinney (Ed.), *The emergence of language*. Mahwah, NJ: Lawrence Erlbaum.

Bock, R. D., & Zimowski, M. F. (1997). Multiple group IRT. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 433–448). New York: Springer.

Braze, D., Tabour, W., Schankweiler, D. P., & Mencl, W. E. (2007). Speaking up for vocabulary: Reading skill differences in young adults. *Journal of Learning Disabilities*, *40*(3), 226–243.

Clark, E. V. (1993). *The lexicon in acquisition*. Cambridge: Cambridge University Press.

Clark, R., Hutcheson, S., & van Buren, P. (1974). Comprehension and production in language acquisition. *Journal of Linguistics*, *10*(1), 39–54.

Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology*, *83*(2), 264–274.

D'Anna, C., Zechmeister, E., & Hall, J. (1991). Toward a meaningful definition of vocabulary size. *Journal of Reading Behavior*, *23*(1), 109–122. doi:10.1080/10862969109547729

Dunn, D. M., & Dunn, L. M. (2007). *Peabody picture vocabulary test*. London: Pearson.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum.

Fengxiang, Fan. (2010). An asymptotic model for the English hapax/vocabulary ratio. *Computational Linguistics*, *36*(4), 631–637.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D., & Pethik, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., Pethick, S., & Reilly, J. S. (1993). *The MacArthur communicative development inventories: User's guide and technical manual*. Baltimore, MD: Paul H. Brokes.

Fleischer, W., Barz, I., & Schröder, M. (2012). *Wortbildung der deutschen Gegenwartssprache* (4. Aufl.). *De Gruyter Studium*. Berlin: De Gruyter.

Gathercole, S., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, *28*, 200–13.

Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language*, *35*(3), 515–531. doi:10.1017/S0305000907008641

Goulden, R., Nation, P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, *11*(4), 341–363.

Grimm, H., & Doil, H. (2005). *Elternfragebögen für die Früherkennung von Risikokindern (ELFRA)* (2nd ed.). Göttingen: Hogrefe.

Hayes, D. P., & Ahrens, M. G. (1988). Vocabulary simplification for children: A special case of 'motherese'? *Journal of Child Language*, *15*(02), 395. doi:10.1017/S0305000900012411

Hoff, E. (2014). *Language development* (5th ed., international ed.). Belmont, CA: Wadsworth Cengage Learning.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236–248.

Jenkins, J. R., Stein, M. L., & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal*, *21*(4), 767–787. doi:10.3102/00028312021004767

Jurish, B., & Würzner, K.-M. (2013): Word and sentence tokenization with hidden Markov models. *JLCL*, *28*(2), 61–83.

Kauschke, C., & Hofmeister, C. (2002). Early lexical development in German. A study on vocabulary growth and vocabulary composition during the second and third year of life. *Journal of Child Language*, *29*(4), 735–757.

Kornai, A. (2002). How many words are there? *Glottometrics*, *4*, 61–86.

Landi, N. (2010). An examination of the relationship between reading comprehension, higher-level and lower-level reading sub-skills in adults. *Reading and Writing*, *23*(6), 701–717. doi:10.1007/s11145–009–9180-z

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343. doi:10.3758/s13428–011–0146–0

Lorge, I., & Chall, J. (1963). Estimating the size of vocabularies of children and adults: An analysis of methodological issues. *The Journal of Experimental Education*, *32*(2), 147–157.

Mochida, A., & Harrington, M. (2006). The Yes/No test as a measure of receptive vocabulary knowledge. *Language Testing*, *23*(1), 73–98. doi:10.1191/0265532206lt321oa

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology*, *40*(5), 665–681. doi:10.1037/0012–1649.40.5.665

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, *20*(2), 233–253.

Naigles, L. R., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs?: Effect of input frequency and structure on children's early verb use. *Journal of Child Language*, *25*, 95–120.

Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, *31*(1), 9–13.

Nation, P. (1993a). Using dictionaries to estimate vocabulary size: Essential, but rarely followed, procedures. *Language Testing*, *10*(1), 27–40. doi:10.1177/026553229301000102

Nation, P. (1993b). Vocabulary size, growth, and use. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 115–135). Amsterdam: John Benjamins.

Nation, P. (2012). Measuring vocabulary size in an uncommonly taught language. *International Conference on Language Proficiency Testing in the Less Commonly Taught Languages*, Bangkok, Thailand,

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, *98*(3), 554–566. doi:10.1037/0022–0663.98.3.554

Perfetti, C., & Hart, L. (2002). The lexical quality hypothesis. *Precursors of Functional Literacy*, *11*, 67–86.

Pregel, D., & Rickheit, G. (1987). *Der Wortschatz im Grundschulalter: Häufigkeitswörterbuch zum verbalen, substantivischen und adjektivischen Wortgebrauch*. Hildesheim: Georg Olms Verlag.

R Core Team. (2015). R: A language and environment for statistical computing: R foundation for statistical computing. Retrieved from http://www.R-project.org/

Reif, M. (2014). PP. Retrieved from https://github.com/manuelreif/PP

Rizopoulos, D. (2006). ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software*, *17*(5), 1–25.

Ruoff, A. (1981). *Häufigkeitswörterbuch gesprochener Sprache: Gesondert nach Wortarten, alphabethisch, rückläufig alphabetisch und nach Häufigkeit geordnet*. Berlin: Walter de Gruyter.

Schmid, H., Fitschen, A., & Heid, U. (2004). SMOR: A German computational morphology covering derivation, composition, and inflection. *Proceedings of the IVth International Conference on Language Resources and Evaluation (LREC 2004)*, 1263–1266.

Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A., & Kliegl, R. (2015). childLex: A lexical database of German read by children. *Behavior Research Methods*. Advance online publication. doi:10.3758/s13428–014–0528–1

Seashore, R. H., & Eckerson, L. D. (1940). The measurement of individual differences in general English vocabularies. *The Journal of Educational Psychology*, *31*(1), 14–38.

Segbers, J., & Schroeder, S. (forthcoming). A vocabulary size test for primary school children in German.

Smith, M. K. (1941). Measurement of the size of general English vocabulary through the elementary grades and high school. *Genetic Psychology Monographs*, *24*, 311–345.

Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading*, *10*(4), 381–398. doi:10.1207/s1532799xssr1004_3

Topping, K. (2015). *What kids are reading: The book-reading habits of students in British schools 2015*. London: Renaissance Learning.

Williams, K. T. (2007). *Expressive Vocabulary Test*, 2nd ed.: *EVT-2*. London: Pearson.