British Journal of Developmental Psychology / Volume 30, Issue 3 / p. 432-445

Effects of morphological Family Size for young readers

Kors Perdijk, Robert Schreuder, R. Harald Baayen, Ludo Verhoeven,

First published: 10 August 2011
https://doi.org/10.1111/j.2044-835X.2011.02053.x

Abstract

Dutch children, from the second and fourth grade of primary school, were each given a visual lexical decision test on 210 Dutch monomorphemic words. After removing words not recognized by a majority of the younger group, (lexical) decisions were analysed by mixed-model regression methods to see whether morphological Family Size influenced decision times over and above several other covariates. The effect of morphological Family Size on decision time was mixed: larger families led to significantly faster decision times for the second graders but not for the fourth graders. Since facilitative effects on decision times had been found for adults, we offer a developmental account to explain the absence of an effect of Family Size on decision times for fourth graders.

Background

According to Miller (1991, p. 238), a child learns about 3,750 words each year, 10 new words every day. Miller emphasizes how amazing this achievement is: 'Clearly, a learning process of great complexity goes on at an impressive rate in every normal child'. How does a child manage to learn such a great number of words in such a short period of time? Decomposing a novel unfamiliar word into familiar morphemes is probably important for this achievement. This paper focuses on the relation between the morphological structure of words, visual word recognition, and the mental lexicon of children.

For most languages, the majority of the words in the language are morphologically complex. That is, these words consist of more than one morpheme, the smallest unit in a language that has a function or meaning. Morphologically complex words tend to occur with a lower frequency than monomorphemic words. When readers (both adult and young readers) encounter a novel word (which by definition has a very low frequency for that child), it is therefore likely to be morphologically complex (see Baayen, 2001).

In this paper, we will study the growth and changing structure of the mental lexicon for children in primary school, using morphological Family Size. For this purpose, we use visual lexical decision instead of offline tasks (see, e.g., Anglin, 1993 for offline studies on Vocabulary Knowledge). The morphological family of a (monomorphemic) word is defined by Schreuder and Baayen (1997) as the set of all words that contain that particular word as a morpheme. Thus, words such as 'homework', 'worker', and 'workmanship' belong to the morphological family of the word 'work'. The Family Size of the word 'work' is the type count of all words that are morphologically related to 'work'. Family Size has been found to have an effect on lexical decision times (e.g., Baayen, Lieber, & Schreuder, 1997; De Jong, Schreuder, & Baayen, 2000; Schreuder & Baayen, 1997). We will first describe how Family Size effects for adult readers have been explained.

For adult readers, Dutch words with a large Family Size were responded to faster and with fewer errors, than words with a small Family Size in visual lexical decision (Schreuder & Baayen, 1997). These results have been replicated for several languages, among which English and Finnish (e.g., Baayen, Feldman, & Schreuder, 2006; Baayen *et al.* 1997; Hyönä & Pollatsek, 1998). Schreuder and Baayen (1997) explain the Family Size effect using a model of morphological processing called the parallel dual-route model (Schreuder & Baayen, 1995). This model distinguishes between two parallel routes of morphological processing, a full form representation route and a morphological decomposition route. Schreuder and Baayen (1995) argue that identification of a morpheme involves activation of its form-specific access representation (see, e.g., Balota, Yap, & Cortese, 2006), which in turn will activate the lemma node (abstract word representation), and its semantic and syntactic nodes. In their turn, the activated semantic and syntactic nodes will activate related lemmas at the lemma level. For example, the word 'work' will activate the lemma *work*, which in turn will activate the semantic nodes denoting 'work'. The semantic node for 'work' will then activate the nodes denoting 'homework' and 'workaholic'. By spreading activation, these nodes will activate the lemmas *work*, *homework*, and *workaholic*. The more lemmas are activated, the more activation spreads in the mental lexicon. Accumulating activation in the mental lexicon signals a higher probability that the presented word is an existing one (see Grainger, 1990); hence, lexical decision latencies will be faster for words with larger families. Interestingly, the computational model of morphological processing proposed by Reichle and Perfetti (2003) can simulate effects of morphological Family Size on word identification.

It is not self-evident that we would find an effect of Family Size for young readers as well. From the reasoning by Schreuder and Baayen (1997), summarized above, it follows that the Family Size effect reflects the presence of stored and connected morphologically related family members in the mental lexicon. Vocabularies of children are smaller in size compared to adults, and there will be fewer connections between lexical representations in their mental lexicon. Children may not have a sufficient number of stored and connected family members in their mental lexicon, for a Family Size effect to occur. Nevertheless, if the explanation for the adult data, postulated by Schreuder and Baayen (1997), holds, effects of Family Size will emerge at some point in lexical development. Family Size effects, therefore, allow us to study the growth and change in organization of the mental lexicon for beginning readers by means of providing evidence for storage and connectedness of lexical representations.

We will first review the evidence on morphological effects for children reported in the literature. Effects of Family Size have been reported in a compound explanation task for 4-year-old children (Krott & Nicoladis, 2005; Nicoladis & Krott, 2007). However, these studies do not say anything about how *written* words are represented, because 4-year-old children do not have written word representations in their mental lexicon yet. Carlisle and Katz (2006) reported effects of Family Size on the accuracy of word reading for 10-year-olds and 12-year-olds. However, reading aloud involves early stages of word recognition, while lexical decision also involves later, more semantic processing (see, e.g., Balota *et al.*, 2006). It is known that children are aware of morphology early in language development (see, e.g., Bloom, 2000). This does not necessarily imply that word morphology is involved in word recognition and that young children's mental lexicons are organized along morphological lines. Thus, the above studies do not tell us anything about effects of morphology on children's lexicons in later stages of word reading. The present study precisely focuses on the role of word morphology in the lexicons of children and on subsequent effects on later stages of word reading. The results of the present study can therefore provide important constraints for developmental models of the mental lexicon and word reading.

Using familiar, monomorphemic words with varying family sizes enables us to test whether other complex words of the same morphological family are stored. While children gradually learn more morphologically complex family members and store them as units in their mental lexicons, sharing connections with other family members, larger effects of Family Size will occur. In this way, we can study vocabulary growth between different grades with the same set of monomorphemic words. This approach differs considerably from previous approaches, using primarily offline, meta-linguistic tasks for studying structure and growth of the mental lexicon These methods are prone to large errors of estimation (see, e.g., Anglin, 1993), which will be smaller in an online test of Vocabulary Knowledge, as is proposed in the present study.

In the present study, we explore effects of Family Size for second grade and fourth grade readers, using a visual lexical decision task. Fourth grade readers will have larger vocabularies than second grade readers. If newly learned words become connected to previously acquired words along morphological lines, we would predict larger Family Size effects for older children.

Because of the empirical evidence indicating that Family Size effects are partly semantically driven (e.g., De Jong *et al*., 2000), one should gauge to what extent effects of Family Size are related to other semantic measures influencing word recognition. Two of these semantic measures, for which effects on word recognition have been studied extensively, are Age Of Acquisition and Imageability (see, e.g., Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Brysbaert & Ghyselinck, 2006; De Jong, 2002; Juhasz, 2005; Menenti & Burani, 2007). Therefore, we included these variables to disentangle their independent contributions to reaction times in lexical decision. Note that Family Size and Age Of Acquisition have been shown to have independent effects on response latencies of lexical decision in adults (De Jong, 2002, Ch. 6).

Method

Participants

Two groups of Dutch primary school children participated in this study, 59 children (35 boys and 24 girls) from the second grade (mean age = 8.1 years, standard deviation [*SD*]= 5.4 months) and 59 children (30 boys and 29 girls) from the fourth grade (mean age = 10.1 years, *SD*= 6.1 months). All participants were monolingual speakers of Dutch from lower middle income families. Schools were situated in the same Dutch city and used the same reading method. For every school, children from both grade levels participated (therefore the two cohorts were balanced with respect to socioeconomic status). For all participants, recent Decoding Skill scores were available, assessed by a standardized reading test (Verhoeven, 1995). This test measures how well children are able to decode written words. The mean Decoding Skill was 44.7 (*SD*= 21.2) words per minute for second grade children and 73.2 (*SD*= 18.4) words per minute for fourth grade children. Both groups on average have scores in the third quartile compared to the norm population (the third quartile comprises scores between 37 and 53 for second grade and scores between 67 and 78 for fourth grade). Furthermore, Vocabulary Knowledge scores were also available for all children, assessed by a standardized test. Different tests were used for second (Verhoeven & Vermeer, 1993) and fourth (Verhoeven, 1999) grade children. The mean vocabulary score was 42.3 (*SD*= 4.7) for second grade children and 96.9 (*SD*= 9.2) for fourth grade children. These scores also were in the third quartile compared to the norm population, just below the national mean (the third quartile comprises scores between 40 and 43 for second grade and scores between 92 and 100 for fourth grade).

Materials

The stimuli used in the present study (see the Appendix) were taken from a previous study with adult participants (Perdijk, Schreuder, & Verhoeven, 2005). The material consisted of 113 nouns and 97 verbs. All words are listed in the sub-corpus for 4- and 5-year-olds in the corpus collected by Schrooten and Vermeer (1994) which is derived from spoken and written materials for children. This particular sub-corpus covers the lowest age range (4- and 5-year-olds) available in this corpus. An additional inclusion criterion was being listed in at least two of the three other grade-specific sub-corpora (6- and 7-year-olds, 8- and 9-year-olds, and 10- and 11-year-olds). These inclusion criteria were used to find target words that most of the children in the study would know. The mean Lemma Frequency of the final set of target words was 1,507 (*SD*= 2,335, range = 130–1,9597) per 42.2 million in an adult corpus (CELEX; Baayen, Piepenbrock, & Gulikers, 1995; a corpus containing 42.2 million Dutch words mainly from newspaper articles, with several word-related measures attached to them). The Family Size counts of the target words (from CELEX) were adjusted in two ways. First, we excluded all semantically opaque family members from the Family Size count.[1] Second, because our aim was to trace the change in effect of Family Size with increasing age (and thus vocabulary size), we only excluded extremely low frequency family members, with a frequency lower than 7 per 42 million in CELEX. Using this criterion, we include many family members that are not yet known by the younger children. However, in this way, we can capture the effects of an increasing vocabulary size. The final set of 210 target items had a mean Family Size of 4.81 (*SD*= 6.14, range = 0–44). Furthermore, for each word, we obtained the Imageability ratings from Van Loon-Vervoorn (1985). For these ratings, participants indicate on a 7-point scale how easily words can be depicted. For our stimuli, this resulted in a mean Word Imageability of 5.60 (*SD*= 1.09, range = 2.00–6.87). We collected the Bigram Frequency of each word from CELEX, consisting of the mean log frequency of all two-letter strings in a word. The mean Bigram Frequency was 8.9 (*SD*= 0.58, range = 7.3–10.3). We also collected the Neighbourhood Size of the words from CELEX, consisting of the number of words resembling the stimulus word orthographically. More precisely, a word was considered a neighbour of the stimulus word when it differed at no more than one position and was maximally one letter longer than the stimulus word. The mean Neighbourhood Size was 28.3 words (*SD*= 33.0, range = 0–252). Age Of Acquisition ratings for the words was obtained from Ghyselinck, De Moor, and Brysbaert (2000). For the words for which these ratings were not available, we carried out a rating ourselves in the same way. We asked 48 participants to indicate the age at which they think they have learned these words. The mean Age Of Acquisition for our stimuli was 5.7 (*SD*= 1.3, range = 2.5–10.4).

In order to create as many non-word items (requiring a no-response) as word items (requiring a yes-response), for each target word, a corresponding pseudoword was constructed. Each target word was altered in one or two letter positions to create a word-like, phonotactically legal, pseudoword. The results for these pseudowords were not analysed but merely served as fillers for the lexical decision task. Four experimental lists were constructed. All target words and pseudowords appeared once in each list. Two random orders were generated for all items and these orders were reversed to get the two additional sequences. An item at the second position in one of the random orders was at pre-final position in the reversed sequence. Therefore, on average, every item was presented halfway the experiment.

Procedure

Two children were tested in parallel during each session. In each session, both participants individually performed a lexical decision task on a computer. Because the number of trials was rather large, trials were divided over two sessions. At the beginning of a session, participants were instructed to decide as quickly and accurately as possible whether a letter string, appearing on the computer screen, was an existing Dutch word.

Each session started with 20 practice trials. Each word was preceded by a fixation mark (lasting for 750 ms) in the middle of the screen. After a 500 ms blank screen, the stimulus appeared at the same position as the fixation mark. The maximal presentation duration of the stimulus was 4,000 ms. Responses and reaction times were recorded using a button box. The left button was used for a no-response and the right button was used for a yes-response. This was reversed for left-handed children. A response by the participant terminated the trial immediately. A response resulted in a 1,000 ms blank screen, followed by a new trial. There were two breaks during each session for the second graders and one break for the fourth graders. A session lasted approximately 30 min for the second graders and 20 min for the fourth graders.

Results

Data reduction

First, we removed all trials on which participants responded faster than 250 ms or slower than 4,000 ms (fewer than 1% of the data). For the remaining trials, items were removed on the basis of the error rates for the particular age groups. The second graders made more than 50% errors on 26 items, which were subsequently removed from the data for this grade (see Appendix). The fourth graders had error rates higher than 50% on five items, which were subsequently removed from the data of this grade (see Appendix).

Regression analysis complete data set

As covariates, we included for each item the semantic variables Imageability and Age of Acquisition, as explained in the Introduction. Furthermore, we included for each item three orthographic variables (Length in Letters, Bigram Frequency, and Neighbourhood Size) as covariates to prevent misinterpreting orthographic effects as being semantic effects. Finally, we also included Lemma Frequency for each item since this variable is known to be an important predictor in lexical decision. To reduce skewness of the distributions of the covariates, Family

Size, Lemma Frequency, and Neighbourhood Size were log transformed. For the measure of Family Size, this transformation at the same time reduces the risk of overestimating the family sizes of our materials. Besides these word-related covariates, we also included a participant-related covariate, Decoding Skill, in our model for each participant and we added the factor Grade, indicating whether the child was in second or fourth grade.

We fitted a mixed model of covariance for all children's decision time data (decision times were log transformed to reduce skewness of the distribution of the decision times). The model was fitted for the 184 words for which the second graders had more than 50% correct. To prevent problems caused by introducing undesired collinearity in our models, the different word-related covariates (Family Size, Lemma Frequency, Word Imageability, Age of Acquisition, Length in Letters, Bigram Frequency, and Neighbourhood Size) were residualized first. The residualization process involved predicting each predictor from all of the other word-related predictors, using linear regression models following Baayen *et al.* (2006). The unexplained residuals of each predictor remaining after this procedure were added to the data set. These residualized word-related predictors, the participant-related predictor Decoding Skill, and the factor Grade (second, fourth) were entered into the model for all children. To investigate the main question, whether children at different ages are differently sensitive to the effect of Family Size, we also included an interaction term of Grade with Family Size in the regression equation. No other interactions were investigated, following Harrell (2001, p. 33) who advises to only include interactions of numerical covariates when there is theoretical reason to do so. Inspection of the residuals of our model for response latencies of the combined data set revealed marked non-normality. We therefore removed outliers with standardized residuals outside the interval [−2:5; 2:5] (2.04% of the data set), and re-fitted our model (see, e.g., Crawley, 2002). The residuals of this trimmed model were approximately normally distributed, indicating that removal of influential outliers resulted in a model with better goodness of fit. We included all variables in the final model that were at least marginally significant. The results of our final model are presented in Table 1 (upper part). To increase readability, we do not report statistics on each predictor in the running text.

Table 1. Mixed models of covariance for the combined data set

Covariate	$\hat{\beta}$	***t*-value, *df*= 17,178**	***p*-value**
Response latencies			
Family Size	−0.0076	−2.08	.0374
Grade	−0.1032	−2.17	.0298
Grade × Family Size	0.0125	1.83	.0674

Covariate	$\hat{\beta}$	*t*-value, *df*= 17,178	*p*-value
Lemma Frequency	−0.0188	−6.86	<.0001
Decoding Skill	−0.0076	−7.78	<.0001
Imageability	−0.0218	−3.93	.0001
Length in Letters	0.0421	7.21	<.0001
Age of Acquisition	0.0198	4.10	<.0001
Bigram Frequency	0.0567	5.56	<.0001
Neighbourhood Size	−0.0322	−6.40	<.0001
Covariate	$\hat{\beta}$	*z*-value, *df*= 21070	*p*-value
Accuracy			
Family Size	0.1415	2.36	.0138

Note. R^2 of the fitted response latency model is .49; Somers' Dxy of the fitted accuracy model is .59.

The final model shows an effect of Family Size. Words with larger families were responded to faster than words with smaller families. Grade also had an effect on response. The fourth graders had faster decision times (mean = 1,258 ms, *SD*= 573) than the second graders (mean = 1,671 ms, *SD*= 752). To test whether children from the two grades behave differently with respect to Family Size, we tested whether Grade interacted with Family Size. The interaction term for Grade by Family Size is marginally significant. Therefore, we decided to fit separate models for the second and fourth grades (see below). Finally, we looked into the effects of the other covariates. Lemma Frequency, Decoding Skill, Imageability, and Neighbourhood Size showed a facilitating effect on lexical decision times. Decision latencies increased with increasing Length in Letters, Age of Acquisition, and Bigram Frequency.

After exploring the response latency data, we report analyses for the accuracy data. This was only done to look for a potential speed accuracy trade-off. Again, we fitted a mixed model of covariance but now for the accuracy data. For this purpose, both correct and incorrect responses on the target words were included in the data set. Successes were coded as ones, and failures as zeros. Success or failure on a trial was the binomial-dependent variable. In the accuracy analyses presented in Table 1 (lower part), a minus sign preceding the $\hat{\beta}$ indicates more error proneness and a positive value indicates less error proneness with an increase of the value of the covariate.

For the accuracy model (Table 1, lower part) of the combined data set, the interaction of Grade with Family Size did not contribute even marginally significantly to the prediction of accuracy and was therefore not included in the final model. The final model again shows an effect of Family Size, indicating that words with a larger family elicited fewer errors. The data also show an effect of Grade. Fourth graders (mean error rate = 11%) reacted more accurately than second graders (mean error rate = 23%). An increase in Lemma Frequency, Decoding Skill, Word Imageability, and Neighbourhood Size led to a decrease in the error rate. In contrast, an increase in Length in Letters, Age of Acquisition, and Bigram Frequency caused an increase in error rate. Note that there is thus no evidence for a speed accuracy trade-off for any of the variables. Given that the dependent variable of the accuracy model was binomial, R^2 could not be computed. We therefore report in Table 1 another goodness of fit variable for our model, *Somers' Dxy rank correlation* (Somers, 1962). This measure of goodness of fit gives a value between 1 and –1. A value of 1 indicates a perfect match between observed response and prediction by the model, –1 indicates a perfect mismatch.

To explore the interaction between Family Size and Grade in the decision latency model further, we also fitted models for the response latency data for the second and fourth grade children separately.[2]

The same fitting procedures were used as for the combined data, except for adding the participant-related covariate Vocabulary Knowledge to the models (and excluding the factor Grade). This covariate could not be included in the model for the complete data set, since the tests used were different for the second and fourth grade children.

Regression analysis: second grade children

All covariates contributed at least marginally significantly to the explanatory power of our statistical model. Inspection of the residuals of our model for decision time of the combined data set revealed marked non-normality. We therefore removed outliers with standardized residuals outside the interval [–2.5; 2.5] (2.20% of the data set), and re-fitted our model (see, e.g., Crawley, 2002). The residuals of this trimmed model were approximately normally distributed, indicating that removal of influential outliers resulted in a model with better goodness of fit. The results of our final model are presented in Table 2. Family Size shows a facilitative effect on decision times. Additionally, decision times decrease with increasing Lemma Frequency, Decoding Skill, Word Imageability, and Neighbourhood Size. An increase in Length in Letters, as well as an increase in Age of Acquisition, Bigram Frequency, and Vocabulary Knowledge leads to an increase of the decision times.

Table 2. Mixed models of covariance for second graders

Covariate	$\hat{\beta}$	*t*-value, *df*= 7,739	*p*-value
Response latencies			
Family Size	−0.0198	−2.12	.0344
Lemma Frequency	−0.0461	−5.92	<.0001
Decoding Skill	−0.0087	−6.19	<.0001
Imageability	−0.0323	−5.07	<.0001
Length in Letters	0.524	7.81	<.0001
Age of Acquisition	0.0196	3.56	.0004
Bigram Frequency	0.0661	5.67	<.0001
Neighbourhood Size	−0.0483	−8.37	<.0001
Vocabulary Knowledge	0.0162	2.55	.0109

Note. R^2 of the fitted response latency model is .45.

Regression analysis: fourth grade children

The same procedure for the second graders was also used for a regression model of decision times for the fourth graders. The main difference with the model for decision times for the second graders is the absence of an effect of Family Size (*p*= .3112) in the model for fourth graders. Furthermore, Vocabulary Knowledge also did not have a significant effect on decision times. Again, on the basis of the residuals, influential outliers were identified and were removed from our model (2.10%), after which the model was re-fitted (see Table 3). For the final model, an increase of Lemma Frequency, Decoding Skill, Word Imageability, and Neighbourhood Size is associated with shorter decision times. An increase of Length in Letters, Age of Acquisition, and Bigram Frequency leads to longer decision times.

Table 3. Mixed models of covariance for fourth graders

Covariate	$\hat{\beta}$	*t*-value, *df*= 10273	*p*-value
Response Latencies			
Lemma Frequency	−0.0470	−6.61	<.0001

Covariate	$\hat{\beta}$	*t*-value, *df*= 10273	*p*-value
Decoding Skill	−0.0072	−5.30	<.0001
Imageability	−0.0097	−1.70	.0884
Length in Letters	0.0363	5.96	<.0001
Age of Acquisition	0.0258	5.27	<.0001
Bigram Frequency	0.0511	5.02	<.0001
Neighbourhood size	−0.0328	−4.78	<.0001

Note. R^2 of the fitted response latency model is .37.

In summary, we find an effect of Family Size on lexical decision times for the second graders only, but not for the fourth graders.

Discussion

In the present study, we investigated the structure and growth of the mental lexicon for young readers, using effects of morphological Family Size as our yardstick. In this way, we investigated the role of morphology in the fast growing mental lexicon of children. Previous research provided evidence for effects of morphological Family Size in the adult mental lexicon (e.g., Baayen *et al*., 1997; De Jong *et al*., 2000; Schreuder & Baayen, 1997). Since studies with young readers might provide insight on the structure and growth of the mental lexicon during literacy development, we investigated the independent contribution, apart from other covariates, of morphological Family Size on word recognition in two age groups. One would expect that older children, having a larger vocabulary, would show a stronger effect of Family Size.

Separate, group-wise analyses showed the following pattern with respect to morphological Family Size: there was a facilitative effect on decision times for the younger, second grade children but there was no effect of Family Size on decision times for the older, fourth grade children. Fourth grade children have larger vocabularies and morphological families that are at least as large as those of the second grade children. However, this is not reflected in our results. The opposite seems to be true, the second grade children show more evidence for an organization of the mental lexicon along morpheme-based families than the fourth graders. We will return to this issue later.

The other covariates related to word meaning, Imageability and Age Of Acquisition, showed their own independent influence on word recognition in the analyses. This is in line with previous work on adults by De Jong (2002, Ch. 6).

Furthermore, for both grades, facilitative effects of word frequency, Neighbourhood Size, and Decoding Skill were obtained on decision times. Note that the effect of written word frequency, which is an established and very stable effect in psycholinguistics, is already present for second graders, similar to results obtained by Burani, Marcolini, and Stella (2002). Longer words and higher Bigram Frequency slowed down decision times for both grades. Vocabulary Knowledge only exerted an effect for the children in second grade. Children with better Vocabulary Knowledge were slower when making lexical decisions. The stability of the effects of almost all predictors in our models indicates that the absence of the Family Size effects for fourth graders on decision times is not just a result of the fourth graders performing the lexical decision task in a different manner.

It is remarkable that the second graders show an effect of Family Size on decision times, whereas the fourth graders do not. Perdijk *et al.*, (2005) reported an effect of Family Size on decision times for an adult population using the same materials as in the present study, indicating that the effect is not simply decreasing with age. Moreover, one would expect older children to have a larger vocabulary, and therefore, larger word families. Note that our study is not the only study indicating stronger effects of Family Size for younger children than for older children. Carlisle and Katz (2006) report an interaction between Family Size and age on the accuracy of word naming, but they fail to interpret this interaction. Inspection of the condition means of their study indicates that the effect of Family Size is relatively larger for the 10-year-old children than for the 12-year-old children. The study of Carlisle and Katz (2006) differs in many aspects from the present study, such as the fact that they use a naming task that involves less semantic processing. In contrast to our study, they do find an effect of Family Size for 10-year-old children. Nonetheless, both studies indicate a decline of effects of Family Size over time. Summarizing, the numerous studies (e.g., Baayen *et al.*, 1997; Baayen *et al.*, 2006; Hyönä & Pollatsek, 1998; Perdijk *et al.*, 2005; Schreuder & Baayen, 1997) reporting a Family Size effect for adults, and the results reported by Carlisle and Katz (2006) and the results of the present study, for children, suggest the presence of a Family Size effect early in lexical development and in adulthood. However, the effect on decision times is absent between these two stages, suggesting a U-shaped development.

One tentative way to interpret this U-shaped development is in terms of a reorganization of semantic representations, as a result of a large increase in vocabulary size. As mentioned before, storage of family members in the mental lexicon is not the only assumption that has to be met for a Family Size effect to occur (Schreuder & Baayen, 1997). The stored word representations should also share connections with morphologically related words in the mental lexicon. When many new words have to be stored in the lexicon, it might be the case that existing connections in the mental lexicon should be 'revisited'. This reorganization of the mental lexicon might cause the Family Size effect to be absent temporarily due to this reorganization process. We will illustrate the plausibility of this reorganization process in the following.

Verhoeven and Vermeer (1996, p. 89) report that Dutch children show an increase in vocabulary growth, starting at the end of second grade. They claim that children will know about 6,000 words at the end of second grade and that, from then on, a child's vocabulary will increase with approximately 3,000 words a year (see also Anglin, 1993, for English). The onset of this increase in vocabulary growth is probably due to changes in the educational context of the children. In the Dutch educational system, children start subjects such as history, geography, biology, and basic concepts of economy in third grade. These new subjects bring along reading larger texts with a great number of words that are new for the children. All these new word representations have to be stored and connected in the mental lexicon.

These words can be closely related to words already present in the lexicon, both in meaning and in morphology. Words related to economics are, for instance, 'workaholic' and 'steelworker'. They both share the morpheme 'work' and the meaning of this constituent is the same as the meaning of the monomorphemic word 'work'. Children can construct the meaning of 'steelworker' by morphological and semantic analysis. These words can therefore be connected to the other words present in the mental lexicon that contain 'work' as constituent morpheme.

One can imagine instances in which such an approach could only succeed partially. For example, words that have just been learned may add novel shades of meaning to their constituents. Children probably first know the word 'earth' in its meaning that deals with soil and will learn during geography class that 'earth' is also used to refer to our planet. Despite the discrepancy between these two definitions, they are related. The planet earth consists of soil. The difference in meaning triggers an adaptation of the overall semantic properties of the word 'earth'. The network of connections within the family of 'earth' (for instance 'earthquake') should therefore be revisited and reorganized. This reorganization might alter spreading activation between family members and, hence, lead to the absence of a Family Size effect on the decision times for the fourth grade children.

If the above hypothesis on reorganization of semantic relations in the mental lexicon holds, one would expect it to be applicable to other age groups and/or situations as well. For example,[3] adults who are engaged in advanced studies might also experience a reorganization of the semantic relations between words related to the domain of their studies, resulting in an absence of Family Size effects for these kinds of words. This possibility could be explored in future research.

An alternative explanation[4] for the differential effect of Family Size is the following. Younger children (second graders) might rely more on phonological recoding of written words, as is suggested by previous studies (see, e.g., Harm & Seidenberg, 2004), than older children (fourth graders). Therefore, the effect of Family Size on decision times for second grade children might arise from connections between phonological representations of words within the spoken lexicon, rather than from connections between orthographic representations within the written lexicon. If fourth graders do not use the intermediate step of recoding words phonologically and have not yet acquired morphological connections between the orthographic representations in their written lexicon, the Family Size effect would be absent.

In sum, the present study reports evidence suggesting that the mental lexicon is organized along morpheme-based families even for young readers (second graders). Furthermore, the results suggest a non-linear development of the effect of morphological Family Size, possibly triggered by reorganizing the semantic connections between morphologically related words in the mental lexicon. A second explanation could be that the effect on decision times for the second grade children might be due to connections in the spoken lexicon (because they still use phonological recoding as a reading strategy). In contrast, the absence of such an effect for fourth graders could be caused by an absence of these connections in their written lexicon.

Footnotes

- 1 *For example for a word as 'brand' (fire), opaque words such as 'branding' (surf) were excluded.*
- 2 *Additional group-wise analyses of the accuracy data did not reveal speed accuracy trade-offs.*
- 3 *We thank an anonymous reviewer for this suggestion.*
- 4 *We thank Cristina Burani for this suggestion.*

Appendix

Avontuur (adventure), baden (to bath), balkon (balcony), band (bond), barsten (to burst), beek (stream), beker (cup), beton (concrete), beven (to shake), bieden (to bid), blaffen (to bark), blijken* (to appear), blinken (to shine), bocht (curve), bonzen (to pound), boodschap (message), bord (plate), broek (trousers), broer (brother), brok (lump), brullen (to roar), buit (catch), bukken (to bend), cent (cent), dal* (valley), deken (blanket), denderen* (to hurtle), douche (shower), dreunen* (to boom), durven (to dare), dwingen (to force), echo (echo), emmer (bucket), etiket[#] (label), flitsen (to lighten), fluisteren (to whisper), fronsen* (to frown), fruit (fruit), gapen (to yawn), garage (garage), giechelen (to gigle), gieten (to pour), gillen (to scream), glinsteren (to sparkle), gloed* (glow), gloeien (to glow), gluren (to peek), graan (grain), grap (joke), grijns (grin), grijnzen (to grin), grinniken (to chuckle), groeten (to greet), grommen (to growl), grot (cave), hijgen (to puff), hijsen (to heave), hinderen* (to hinder), horizon (horizon), horloge (watch), insekt (insect), jagen (to hunt), juichen (to cheer), kanon (canon), karton (cardboard), karwei (chore), kerel (lad), ketel (kettle), kilo (kilo-gram), kin (chin), klagen (to complain), klant (customer), knagen (to gnaw), kneden, (to knead), knielen (to kneel), knipperen (to blink), knol (tuber), koffer (suitcase), koning (king), koren (corn), korst (crust), kraag (collar), kreunen (to moan), krijsen (to screech), krimpen (to shrink), kronkelen* (to wind), kudde (herd), kus (kiss), ladder (stepladder), laden (to load), lawaai (noise), leunen (to lean), liegen (to lie), lijden (to suffer), liter (litre), loeren (to glare), lucifer* (matchstick), manier (manner), matras (mattress), matroos (sailor), melden (to report), mengen (to blend), metselen (to lay bricks), minuut (minute), mompelen (to mumble), mouw (sleeve), mus (sparrow), muts (hat), neef (cousin), ober (waiter), oceaan* (ocean), oever (shore), oom (uncle), paradijs (paradise), parfum (perfume), park (park), peinzen* (to ponder), pet (cap), piekeren* (to brood), piepen (to squeak), pijl (arrow), piloot (pilot), plafond (ceiling), plein (square), plek (place), plezier (pleasure), plukken (to pick), podium* (stage), poes (cat), poos (while), proberen (to try), raket (rocket), razen* (to rage), rekenen (to calculate), rest (rest), rij (row), rijst (rice), rillen (to shiver), rinkelen (to tinkle), romp# (torso), ruiken (to smell), ruisen# (to rustle), ruzie (fight), schamen (to feel embarrassed), schande (defamation), schelden (to call someone names), scheut# (sprout), schitteren (to shine), schoot (lap), schrapen (to scrape), sissen (to hiss), slenteren* (to saunter), slikken (to swallow), sloot (ditch), smeken (to beg), smijten (to fling), snauwen (to growl), snikken (to sniffle), snor (moustache), snuiven (to sniff), sok (sock), spijten (to regret), spleet (cleft), sprookje (fairytale), spul (stuff), stampen (to stamp one's foot), staren (to stare), stier (bull), stinken (to stink), stoep (sidewalk), stoken (to heat), strekken (to stretch), strelen* (caress), struikelen (to trip), stuiven (to swirl), taart (cake), tante (aunt), teen (toe), temperatuur (temperature), trots (pride), truc* (trick), trui (sweater), turen (to stare), ui (onion), vaas (vase), vacht (fur), verdriet (sadness), verrassen (to surprise), vijver (pond), villa (villa), vork (fork), vuist (fist), waas# (haze), wal* (shore), wang (cheek), wapperen (to flap), weven (weave), wiegen (to cradle), woelen* (to toss and turn), worm (worm), worstelen (wrestle), woud (woods), wringen* (to wrench), wuiven* (to wave), zeuren (to nag), zuigen (to suck), zwaard (sword), zwemmen (to swim), zweven (to hover), zwijgen (to keep silent)

Note. Items marked with an asterisk were excluded from the analyses for the second graders and items marked with a hash were excluded from the analyses for the second graders and the fourth graders.

References

Anglin, J. M. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development*, **58**, 1– 166. doi:10.2307/1166112
Crossref | Web of Science® | Google Scholar

Baayen, R. H. (2001). *Word frequency distributions*. Dordrecht : Kluwer Academic Publishers.
Crossref | Google Scholar

Baayen, R. H., Feldman, L. F., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, **55**, 290– 313. doi: 10.1016/j.jml.2006.03.008
Crossref | Web of Science® | Google Scholar

Baayen, R. H., Lieber, R., & Schreuder, R. (1997). The morphological complexity of simplex nouns. *Linguistics*, **35**, 861– 877. doi:10.1515/ling.1997.35.5.861
Crossref | Web of Science® | Google Scholar

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The celex lexical database[CD-ROM]*. Philadelphia , PA : Linguistic Data Consortium, University of Pennsylvania (Distributor).
Google Scholar

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, **133**, 283– 316. doi: 10.1037/0096-3445.133.2.283
Crossref | PubMed | Web of Science® | Google Scholar

Balota, D. A., Yap, M. J., & Cortese, M. J. (2006). Visual word recognition: The journey from features to meaning (A travel update). In M. Traxler & M. A. Gernsbacher (Eds.). *Handbook of psycholinguistics* (2nd ed.). San Diego , CA : Academic Press.
Crossref | Google Scholar

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge , MA : MIT Press.
Crossref | Google Scholar

Burani, C., Marcolini, S., & Stella, G. (2002). How early does morpholexical reading develop in readers of a shallow orthography? *Brain and Language*, **81**, 568– 586. doi:10.1006/brln.2001.2548
Crossref | PubMed | Web of Science® | Google Scholar

Brysbaert, M., & Ghyselinck, M. (2006). The effect of age of acquisition: Partly frequency related, partly frequeny independent. *Visual Cognition*, **13**, 992– 1011. doi:10.1080/13506280544000165
Crossref | Web of Science® | Google Scholar

Carlisle, J. F., & Katz, L. A. (2006). Effects of word and morpheme familiarity on reading of derived words. *Reading and Writing*, **19**, 669– 693. doi:10.1007/s11145-005-5766-2
Crossref | Web of Science® | Google Scholar

Crawley, M. J. (2002). *Statistical computing. An introduction to data analysis using S-plus*. Chichester : Wiley.
Google Scholar

De Jong, N. H. (2002). *Morphological families in the mental lexicon*.MPI series in psycholinguistics. Nijmegen : Max Planck Institute of Psycholinguistics.
Google Scholar

De Jong, N. H., Schreuder, R., & Baayen, R. H. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes*, **15**, 329– 365. doi:10.1080/01690960050119625
Crossref | Web of Science® | Google Scholar

Ghyselinck, M., De Moor, W., & Brysbaert, M. (2000). Age-of-acquisition ratings for 2816 Dutch four- and five-letter nouns. *Psychologica Belgica*, **40**, 77– 98.
Web of Science® | Google Scholar

Grainger, J. (1990). Word-frequency and neighborhood frequency-effects in lexical decision and naming. *Journal of Memory and Language*, **29**, 228– 244. doi:10.1016/0749-596X(90)90074-A
Crossref | Web of Science® | Google Scholar

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, **111**, 662– 720. doi: 10.1037/0033-295X.111.3.662
Crossref | PubMed | Web of Science® | Google Scholar

Harrell, F. (2001). *Regression modeling strategies*. Berlin : Springer.
Crossref | Google Scholar

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance*, **24**, 1612– 1627. doi:10.1037//0096-1523.24.6.161
Crossref | CAS | PubMed | Web of Science® | Google Scholar

Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin*, **131**, 684– 712. doi:1037/0033-2909.131.5.684
Crossref | PubMed | Web of Science® | Google Scholar

Krott, A., & Nicoladis, E. (2005). Large constituent families help children parse compounds. *Journal of Child Language*, **32**, 139– 158.
Crossref | PubMed | Web of Science® | Google Scholar

Menenti, L., & Burani, C. (2007). What causes the effect of age-of-acquisition in lexical processing *Quarterly Journal of Experimental Psychology*, **60**, 652– 661. doi:10.1017/S0305000904006622
Crossref | PubMed | Web of Science® | Google Scholar

Miller, G. A. (1991). *The science of words*. New York : Scientific American Library.
Google Scholar

Nicoladis, E., & Krott, A. (2007). Family size and French-speaking children's segmentation of existing compounds. *Language Learning*, **57**, 201– 228. doi:10.1111/j.1467-9922.2007.00407.x
Wiley Online Library | Web of Science® | Google Scholar

Perdijk, K., Schreuder, R., & Verhoeven, L. (2005). The role of morphological family size in word recognition: A developmental perspective. *Written Language and Literacy*, **8**, 45– 59. doi:10.1075/wll.8.2.05per
Crossref | Google Scholar

Reichle, E. D., & Perfetti, C. A. (2003) Morphology in word identification: A word-experience model that accounts for morpheme frequency effects. *Scientific Studies of Reading*, **7**, 219– 238. doi:10.1207/S1532799XSSR0703_2
Crossref | Web of Science® | Google Scholar

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131– 154). Hillsdale , NJ : Lawrence Erlbaum Associates, Inc.
Google Scholar

Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language*, **37**, 118– 139. doi:10.1006/jmla.1997.2510
Crossref | Web of Science® | Google Scholar

Schrooten, W., & Vermeer A. (1994). *Woorden in het basisonderwijs. 15.000 woorden aangeboden aan leerlingen* [Words in primary school. 15,000 words presented to primary school children], Studies in meertaligheid 6. Tilburg: Tilburg University Press.
Google Scholar

Somers, R. H. (1962). A new asymmetric measure of association for ordinal variables. *American Sociological Review*, **27**, 799– 811. doi:10.2307/2090408
Crossref | PubMed | Web of Science® | Google Scholar

Van Loon-Vervoorn, W. A. (1985). *Voorstelbaarheidswaarden van Nederlandse woorden: 4600 substantieven, 1000 verba, 500 adjectieven* [Imageability values of Dutch words: 4600 nouns, 1000 verbs, 500 adjectives]. Lisse : Swets & Zeitlinger b.v.
Google Scholar

Verhoeven, L. (1995). *Drie-Minuten-Toets* [Three-Minutes-Test]. Arnhem: CITO.
Google Scholar

Verhoeven, L. (1999). *Leeswoordenschat* [Reading Vocabulary]. Arnhem: CITO.
Google Scholar

Verhoeven, L., & Vermeer, A. (1993). *Woordenschattoets* [Vocabulary Test]. Arnhem: CITO.
Google Scholar

Verhoeven, L., & Vermeer, A. (1996). *Taalvaardigheid in de bovenbouw* [Language proficiency in grades 3 to 6]. Tilburg : Tilburg University Press.
Google Scholar