Effects of family size for complex words

Raymond Bertram
Department of Psychology,
University of Turku, Turku, Finland

R. Harald Baayen & Robert Schreuder
Interfaculty Research Unit for Language and Speech
University of Nijmegen
The Netherlands

SUBMITTED, PLEASE DO NOT QUOTE

Address all correspondence to:
Raymond Bertram
Department of Psychology
University of Turku
Lemminkäisenkatu 14-18 B
20520 Turku, Finland

Abstract

Schreuder and Baayen (1997) reported that in visual lexical decision response latencies to a simplex noun are shorter when this noun has a large morphological family, i.e., when it appears as a constituent in a large number of derived words and compounds. The present study addresses the question whether the Family Size of the base word of a complex word likewise affects lexical processing. Six experiments are reported that show that Family Size indeed plays a role for both inflected and derived words. Post-hoc analyses show that the effect of Family Size is driven by the semantically transparent family members, and that this effect is further constrained by semantic selection restrictions of the affix in the target word.

KEYWORDS: morphological processing, family size effect, complex words

Frequency of occurrence is recorded in human memory for a wide variety of naturally occurring events, irrespective of age, ability, intentionality, instructions, or practice (Hasher & Zacks, 1984). Within the domain of language processing, the word frequency effect is one of the most robust findings both in speech production (e.g., Jescheniak & Levelt, 1994; Oldfield & Wingfield, 1965) and in language comprehension (e.g., Rubenstein & Pollack, 1963; Whaley, 1978). Words with a high frequency of occurrence are processed faster and more accurately than words with a low frequency of occurrence. Word frequency effects arise not only for monomorphemic words, but also for morphologically complex words (e.g., Baayen, Dijkstra, & Schreuder, 1997; Taft, 1979). In addition, the frequencies of occurrence of the constituents of complex words may also affect lexical processing (e.g., Burani & Caramazza, 1987; Taft, 1979). These frequency effects all concern counts of the number of tokens that instantiate a given word type in language corpora.

Recently, a new frequency component in human memory has been discovered that is not based on frequency of occurrence measured in terms of the number of tokens counted in corpora, but on a type count. Schreuder and Baayen (1997) show that for a monomorphemic noun its morphological Family Size, the type count of the number of derived words and compounds in which that noun occurs as a constituent, co-determines response latencies and response accuracy in visual lexical decision as well as subjective frequency ratings (see also Baayen, Lieber & Schreuder, 1997).[1] Independently of the various token-based frequency counts, words with a large Family Size elicit shorter response latencies, fewer errors, and higher subjective frequency ratings, than words with a small Family Size.

Schreuder and Baayen (1997) did not observe any effect of Family Size in progressive demasking, a task that is primarily sensitive to lexical processing at the early stages of visual identification (Grainger & Segui, 1990). This dissociation suggests that the effect of Family Size arises at more central levels of lexical processing. Consequently, Schreuder and Baayen (1997) explained the effect of Family Size for simplex words by assuming that lexical decisions and subjective frequency ratings are sensitive to the number of activated entries in the central lexicon along the lines of Grainger and Jacobs (1996). When a word like werk, ‘work’, which occurs in a great many morphologically complex words, is read, the entries of its large number of morphological descendants are also activated, leading to a substantial amount of global activation in the central lexicon. This global activation allow participants to execute their responses quickly.

The Family Size effect is remarkable for two reasons. First, it is surprising that complex words that are not presented to participants nevertheless influence their responses. Second, the distinction between inflection and derivation is mirrored in the way in which monomorphemic nouns are processed. The summed token frequencies of the inflectional variants of a monomorphemic noun co-determine response latencies and error scores in visual lexical decision as well as subjective frequency ratings (Baayen, Dijkstra, & Schreuder, 1997; Baayen, Lieber, & Schreuder, 1997; Schreuder & Baayen, 1997). Conversely, the summed token frequencies of the derived words and compounds in which a noun occurs as a constituent do not affect the lexical processing of that noun. Outside the domain of inflection, it is the type count of the number of family members that is the relevant factor (see also De Jong, Schreuder, & Baayen, 1999).

In this study, we use the Family Size effect as a tool to gain further insight into the organization of the central mental lexicon. We do so by investigating whether and under which conditions Family Size effects occur for morphologically complex words. More precisely, we investigate whether the morphological family of a base word (e.g., the family of werk, ’work’), is activated upon visual presentation of werkbaar and werkte, words in which the base werk is embedded.

It is not a-priori self-evident that such an effect will be observed. For instance, it is possible that an effect of Family Size is strictly mediated by the access representation of the simple base form. This would entail that one would not expect a Family Size effect without a co-occurring Base Frequency effect. Models which do not assign an independent role to morphology in word recognition (e.g., Butterworth, 1983; Seidenberg, 1987) would then predict that no Family Size effect should occur.

Alternatively, the Family Size effect might be a more central morphological and semantic effect, as hypothesized by Schreuder and Baayen (1997). If so, it would be irrelevant whether a complex word has been recognized via the access representation of the base. Instead, any complex word standing in a fully regular relation to its base, both formally and semantically, might co-activate its base word at the semantic level. From the base word, activation might then spread further into the morphological family. However, it is not self-evident that all derived words and compounds should reveal an effect of morphological Family Size, as many such formations often have idiosyncratic meanings that are not transparent with respect to the meanings of their constituents.

These considerations lead us to expect effects of Family Size primarily for inflected words. Firstly, inflection does not change the meaning of the base word. Secondly, the occurrence of full-form frequency effects for inflected forms is still a debated issue in the literature (Baayen, Dijkstra, & Schreuder, 1997; Sereno & Jongman, 1997; Taft, 1979; but see Clahsen, Eisenbeiss, & Sonnenstuhl-Henning, 1997). In fact, there are models of lexical processing that regard inflected words as the only complex forms for which on-line decomposition is likely to take place (see, e.g., Caramazza, Laudanna, & Romani, 1988; Niemi, Laine, & Tuominen, 1994). In similar vein, Taft (1994) suggested that inflectional suffixes are automatically stripped off and fed into a syntactic module separate from the lexical system. Even in models in which the distinction between inflection and derivation is not relevant for the early stages of form identification in lexical processing (Schreuder & Baayen, 1995), the distinction between inflection and derivation remains relevant for more central levels of lexical processing and representation. Inflected and derived words have different syntactic and semantic functions. In some linguistic theories, inflection is even assigned to a different component of the grammar than derivation and compounding (see, e.g., Anderson, 1992, and for discussion, Booij, 1993).

However, the boundary between inflection and derivation is a fuzzy one (see, e.g., Booij, 1998). Affixes are situated on a continuum from prototypical inflection to prototypical derivation (Bybee, 1985). In order to study the effect of Family Size for complex words, we have studied affixes that cover the whole range, from prototypical inflection to prototypical derivation. Because the way in which complex words are processed may depend in part on the various different properties of individual affixes (see, e.g., Laudanna & Burani, 1995), we have opted for an experimental approach in which we study the effect of Family Size for one affix at a time. We have selected for our experiments the four affixes studied by Bertram, Baayen, and Schreuder (1999), for two reasons.

First, these affixes represent prototypical inflection (the past-tense suffix -te), inflection on the border of derivation (the comparative suffix -er, which changes the argument structure of the adjective), and prototypical derivation (the derivational suffixes -er and -heid, which, like their English counterparts -er and -ness, form subject nouns and abstract nouns from verbs and adjectives respectively). Second, the role of storage and parsing for these suffixes is well-studied (Bertram et al., 1999), which allows us to ascertain, for instance, to what extent the effect of Family Size for complex words is dependent on on-line parsing.

Experiment 1

Bertram et al. (1999) conducted two experiments on the Dutch past tense suffix -te (e.g., werk-te, 'worked') in which they orthogonally contrasted either high and low Surface Frequency or high and low Base Frequency, while matching for other factors (including Family Size). Their results revealed an effect of Base Frequency (the summed frequency of the inflectional variants of the base) but no effect of Surface Frequency (the frequency of the inflected form itself) for -te. Since verbal inflections in -te stand in a fully regular relation to their bases, we expect a solid effect of Family Size for words with this inflectional suffix as well.

Method

Participants. Sixteen undergraduate students from the University of Nijmegen were paid to participate in the lexical decision experiment. Sixteen different undergraduate students from the same university were paid to perform the subjective frequency rating experiment. All were native speakers of Dutch and had normal or corrected-to-normal vision. The participants in the experiments to follow were from the same population. Each participant was involved in only one of the experiments.

Materials. Forty inflected verb forms with the singular past tense suffix -te were selected from the CELEX lexical database (Baayen, Piepenbrock & Gullekers, 1995), of which twenty had a large Family Size (henceforth Vf^2; see Schreuder & Baayen, 1997) of on average 36 descendants, (range $10-185$) whereas the other twenty had a small Family Size of on average

3 descendants (range 1 – 9) . Since past tense forms cannot occur as constituents of other complex words, the Family Size of the target words themselves (henceforth the full form Family Size) was zero for both conditions. The two sets were matched for Base Frequency[3] (high Vf, 21.0; low Vf, 20.8; all token frequency counts reported are scaled to one million), Surface Frequency (high Vf, 6.3; low Vf 6.7;, geometric mean bigram frequency (high Vf, 13.7; low Vf, 13.8), and word length in letters (high Vf, 6.7; low Vf, 6.5). The materials are listed in the Appendix. An additional 100 filler words were added, consisting of monomorphemic and inflected nouns and verbs, derived words with the deverbal suffix -er, and derived words with the productive de-adjectival suffix -heid. For each of the total of 140 words, a nonword was obtained by changing one to three letters, such that the phonotactics of Dutch were not violated.

Procedure. For the lexical decision task, participants were tested two at the same time in noise-proof experimental booths. They were asked to decide as quickly and accurately as possible whether a letter string appearing on the computer screen was a real Dutch word or not. Each stimulus was preceded by a fixation mark in the middle of the screen for 500 ms. After 50 ms, the stimulus appeared at the same position. Stimuli were presented on Nec Multisync color monitors in white lowercase 18 point Helvetica letters on a dark background and they remained on the screen for 1500 ms. The maximum response time was 2000 ms from stimulus onset. Fifteen practice trials, 8 words and 7 nonwords, were run before the actual experiment. The actual experiment was divided in two blocks of 140 items (70 words and 70 nonwords), and there was a short pause between the blocks. In total, the experiment lasted approximately 20 minutes.

For the subjective frequency ratings, participants were asked to indicate on a seven-point scale how frequently they thought a word was used in current Dutch. In the past, subjective frequency (or familiarity) ratings have been used to check whether corpus-based frequency counts are reliable. However, recent experiments have revealed that subjective frequency ratings are sensitive not only to Surface Frequency, but also to Base Frequency and Family Size, both in Dutch (Schreuder & Baayen, 1997; Bertram et al. 1999) and in English (Baayen, Lieber, & Schreuder, 1997). Even though an off-line task, the subjective frequency rating task apparently taps into various aspects of lexical representation and processing. We use the subjective frequency ratings to obtain independent confirming evidence for the reaction time data.

Results and Discussion

All participants performed the lexical decision task with an overall error rate less than 15%. One item elicited errors at a rate exceeding 30%, and was therefore excluded from the reaction time analyses. All remaining observations were used to calculate the mean response latencies for the different test conditions, as can be found in Table 1. Error scores were calculated over all words in this experiment and all other experiments reported. In addition to the mean error scores, Table 1 also lists the mean subjective frequency scores.

PLACE TABLE 1 APPROXIMATELY HERE

A paired t-test for participants and a standard two-sample t-test for items showed the predicted difference between the two conditions in the expected direction: the verbs with a large Family Size were responded to significantly faster than the verb inflections with a small Family Size ($t1(15) = -2.06, p < .05$; $t2(37) = 1.73, p < .05$). (As frequency effects, if present, lead to faster response latencies, lower error scores, and higher subjective frequency ratings, we have used one-tailed tests throughout this paper.) The words with a low Family Size also elicited higher error rates ($t1(15) = 3.11, p < .005$; $t2(38) = 3.08, p < 0.005$).

For the subjective frequency ratings, we also obtained a statistically significant difference in the by-participant analysis and a marginally significant difference in the by-item analysis ($t1(15) = 6.17, p < .001; t2(37) = 1.62, p < .06$).

We conclude that the size of the morphological family of the base word indeed co-determines processing of inflected verbs in Dutch. Recall that Bertram et al. (1999) observed an effect of Base Frequency and no effect of Surface Frequency for inflected verbs with the suffix -te. Thus, it may be that the effect of Family Size that we observe in this experiment arises due to the mediation of the access representation of the verbal base form.

Experiment 2

In this experiment we progress from a pure inflectional suffix to a suffix that straddles the border between inflection and derivation, the comparative suffix -er. Bertram et al. (1999) did not observe an effect of Base Frequency for this suffix, but only an effect of Surface Frequency. An effect of Family Size for comparative adjectives would imply that the effect does not critically hinge upon the activation of the base form itself.

Method

Participants. Sixteen students participated in the lexical decision experiment. Sixteen different students did the subjective frequency rating.

Materials. Forty comparatives were selected from the CELEX database, of which twenty had a large Family Size (52, range 10 – 171, full form Family Size is zero), whereas the other twenty had a small Family Size (3, range 0 – 9, full form Family Size is zero). The two sets were matched for Base Frequency (high Vf: 86.4; low Vf: 86.5), Surface Frequency (high Vf: 3.0; low Vf: 3.1), geometric mean bigram frequency (13.9 for both conditions), and word length in letters (high Vf: 6.7; low Vf: 7.1). The materials are listed in the Appendix. The additional 100 filler words consisted of monomorphemic and inflected nouns, adjectives and verbs, and derived words with the deverbal suffixes -sel and -ing. For all the 140 words, a nonword was obtained by changing one to three letters, such that the phonotactics of Dutch were not violated.

Procedure. The procedure was identical to that of Experiment 1.

Results and Discussion

The data of all participants were included in the analyses, since they all performed with an overall error rate below 15%. All items elicited error rates below 30%, so that they too could all be included in the analyses. The observations were used to calculate the mean response latencies and error scores for the different test conditions, as can be found in Table 2, which also lists the mean rating scores.

PLACE TABLE 2 APPROXIMATELY HERE

A paired t-test for participants and a standard two-sample t-test for items revealed a reliable difference between the two conditions ($t1(15) = 6.4, p < .001; t2(38) = 3.1, p < .005$). The comparatives with a large Family Size were responded to significantly faster than the comparatives with a small Family Size. The subjective frequency ratings yielded a similar pattern. This difference was significant in the by-participant analysis ($t1(15) = 6.17, p < .001$), and marginally so in the by-item analysis ($t2(38) = 1.3, p = .10$). The error data of the reaction time experiment did not reveal any reliable difference ($t1(15) = -1.16, p = 0.13$, $t2(38) = -1.16, p = 0.13$).

Just as in Experiment 1, we observe a clear effect of Family Size for -er in the reaction time data, and to some extent also an effect in the subjective frequency ratings, even though the base verb does not appear to play a significant role as shown by the absence of a Base Frequency effect for this suffix as reported by Bertram et al. (1999) This indicates that activation of the base is not a necessary condition for the Family Size effect to occur.

Experiment 3

In the next experiment we examine a suffix that is fully derivational in nature, the homonym of the comparative suffix that is used to create nouns from verbs (bakker, 'baker'). Many complex words in derivational -er have additional shades of meaning that are so typical of derived words in general. For instance, a baker is normally interpreted as someone whose occupation it is to bake things, rather than someone who happens to bake a cake every once in a while. With Experiment 3 we investigate whether the transition from inflection to derivation and the concomitant introduction of additional shades of meanings affects the Family Size effect.

Method

Participants. Sixteen students participated in the lexical decision experiment. Sixteen different students did the subjective frequency rating.

Materials. Forty derived nouns with the deverbal suffix -er were selected from the CELEX database, from which twenty had a large Family Size (23, range $9-95$, full form Family Size 1.7), whereas the other twenty had a small Family Size (5, range $2-11$, full form Family Size

0.3). The two sets were matched for Base Frequency (high Vf: 32.2; low Vf: 31.6), Surface Frequency (high Vf: 0.4; low Vf, 0.2), geometric mean bigram frequency (both high and low Vf: 13.9), and word length in letters (high Vf: 6.8; low Vf: 6.4). The materials are listed in the Appendix. An additional 100 filler words were added, consisting of monomorphemic and inflected nouns, adjectives and verbs, and derived words with the deverbal suffixes -sel and -ing. For each of the 140 words, a nonword was obtained by changing one to three letters, such that the phonotactics of Dutch were not violated.

Procedure. The procedure was identical to that of Experiment 1.

Results and Discussion

The data of all participants were included in the analyses, since they all performed with an overall error rate below 15%. Eleven items had error rates higher than 30% and were removed from the reaction time analyses (removal of the error-prone items did not substantially affect the matching between the two sets of words). The differences in the response latencies between words with a high versus a low Family Size are not significant ($t1(15) = 1.7, p > .10; t2(27) = 0.9, p > .10$).

In contrast to the response latencies, the subjective frequency ratings show a significant effect of Family Size ($t1(15) = 8.3, p < .01; t2(38) = 2.63, p < .01$). These results prompted us to have a closer look at the lexical decision data. Interestingly, the unusually high number of error-prone items is found primarily in the low Family Size condition (2 in the high-frequency condition and 9 in the low-frequency condition). A chi-square test on the distribution of these error-prone items shows that they are indeed not randomly distributed over the two conditions ($X^2(1) = 4.51, p < .05$, with continuity correction). The error analysis, using all 40 words, also reveals a significant effect of Family Size: 10.0% for the large Family Size condition, and 26.6% for the small Family Size condition ($t1(15) = -6.1, p < .01; t2(38) = -3.0, p < .025$). Means, error rates, and ratings can be found in Table 3.

PLACE TABLE 3 APPROXIMATELY HERE

Summing up, although the reaction time data do not reveal an effect of Family Size, the error data and the subjective frequency data suggest that Family Size may nevertheless have a role to play for derived nouns in -er. In order to ascertain whether a fully reliable effect of Family Size can be observed in the response latencies to complex nouns in -er as well, we ran a further experiment using a regression design as the possibilities of an orthogonal contrast have already been exhausted by the materials of Experiment 3.

Experiment 4

Method

Participants. Seventeen students participated in the lexical decision experiment.

Materials. One hundred and twenty-two derived nouns with the deverbal suffix -er were selected from the CELEX database. These words comprise all words in -er in Dutch derived from monomorphemic base words that are unambiguously verbs. Thus, a word like arbeider ('worker') was not included due to the existence of the noun arbeid ('work') side by side with the corresponding verb arbeiden. These words varied substantially over the variables Family Size (mean 26.3, standard deviation 34.3, median 11, range 0–160), Surface Frequency (mean 3.86, standard deviation 14.37, median 0.33, range 0–120.4) and Base Frequency (mean 128.6, standard deviation 308.5, median 18.6, range 0–2342.3). The mean length of our target words was 6.4 letters with a standard deviation of 0.78. A total of 148 fillers was included in the Experiment. All filler words were low-frequency complex words with a range of derivational suffixes selected from the frequency range [0 – 0.5] per million. This low frequency range was selected in order to avoid 'surprise' effects for our lowest frequency target words. 270 nonwords were constructed by changing one to three letters such that the phonotactics of Dutch were not violated.

Procedure. The procedure was identical to that of Experiment 1.

Results and Discussion

Fifteen words elicited error rates greater than 30%, and were removed from further analyses. The remaining 106 words with acceptable error rates (including 19 words with zero-frequency) were subjected to a multiple regression analysis with Response Latency as dependent variable and Surface Frequency, Base Frequency and Family Size as independent variables. For each frequency measure, we used a logarithmic transformation ($\log(f + 1)$). We checked for possible collinearity in the data matrix by means of the condition index and the matrix of variance-decomposition proportions (Belsley, Kuh, E. & Welsch, 1980). Given a moderately low condition index $\kappa = 13.15$ and given little evidence in the variance-decomposition matrix for linear dependency between our independent variables, we applied a stepwise linear regression analysis. This analysis resulted in reliable facilitatory main effects for Family Size ($c = -71.7, t = -4.74, p < 0.0001$) and Base Frequency ($c = -24.1, t = -4.34, p < 0.0001$) as well as in a small significant inhibitory interaction between Family Size and Base Frequency ($c = 7.3, t = 3.9, p < 0.001$). Inspection of the studentized residuals and the fitted values did not reveal any systematic pattern, suggesting that our data do not violate the assumptions underlying linear modeling.

The inhibitory interaction of Family Size and Base Frequency is difficult to assign a meaningful interpretation, and may be an artifact of the distributions involved — neither response latencies nor (even logarithmically transformed) word frequency counts are fully normally distributed.

We conclude that the Family Size of the base words of our complex words in -er has some facilitatory role to play in lexical processing. Apparently, the effect of Family Size is not restricted to inflected words, but also extends to derived words, even though these derived words often have their individual shades of meaning.

However, compared to the more inflectional homonym of -er, the Family Size effect of the base is more difficult to establish reliably and independently of Base Frequency, which suggests that perhaps the effect of Family Size is stronger for inflected words than for derived words, all other things being equal.

To test this possibility, we designed an experiment in which inflected and derived words sharing the same base word are directly compared By investigating pairs such as werk-te ('worked') and werk-baar ('workable'), it is possible to compare the effect of high versus low Family Size of the base for inflected and derived words, with identical Family Sizes for both kinds of word formation.

Experiment 5

Method

Participants. Thirty students participated in the lexical decision experiment.

Materials. We selected forty monomorphemic base words from the CELEX lexical database. Twenty of these base words had a large Family Size (36.3, range 10 – 185) and twenty of these base words had a small Family Size (3.6, range 1 – 9). The two sets were matched for the frequency of the base and its inflectional variants (21.1, 20.1 respectively). For each base word, we selected the corresponding regular past tense inflected form in -te for inclusion in the experiment. The mean Surface Frequency of the past tense forms with a large Family Size was 6.1, the mean Surface Frequency of the past tense forms with a small Family Size was 6.3. For exactly the same base words, we also selected non-inflectional complex words, 37 derivations (with a variety of affixes) and 3 compounds. This variety was unavoidable because we attempted to pair the inflected words with existing derived words with not too low frequencies of occurrence. The mean frequency of the derived words with a large Family Size was 2.3, the mean frequency of the derived words with a small Family Size was 2.1. Because derived words tend to be substantially less frequent than their base words and the inflected variants of these base words, we could not match the inflected and derived sets of words on Surface Frequency. We will return to this imbalance below.

The 80 target words were equally devided between two experimental lists such that a given base word occurred only once in a given list, either as an inflected form or as a derived form. Each list contained an equal number of derived and inflected forms. Both lists were likewise balanced with respect to Family Size. To the forty target words in each list, we added 100 filler words (both monomorphemic, inflected, and derived words) as well as 140 pseudo words obtained from existing words by changing one to three letters such that the phonotactics of Dutch were not violated. Finally, each list was preceded by 12 practice items (six words, six pseudo words).

Procedure. The procedure was identical to that of Experiment 1.

Results and Discussion

Eight target words (with eight different base words) elicited more than 30% errors. These words, together with their derivational or inflectional paired variants, were removed from the reaction time analyses. Thus, we were left with 32 pairs of target words, which differed orthogonally with respect to Family Size, and which were still matched with respect to the variable Surface Frequency (inflected forms: 6.3 (high) versus 7.0 (low); derived forms: 2.4 (high) versus 2.4 (low)) and Base Frequency (inflected forms: 21.8 (high) versus 22.1 (low); derived forms: 20.4 (high) versus 21.9 (low)). Mean response latencies and error scores are summarized in Table 4.

PLACE TABLE 4 APPROXIMATELY HERE

Analyses of variance revealed reliable effects of Family Size and Type of Word Formation (Inflection versus Derivation) but no interaction (Family Size: $F1(1,29) = 14.1, \mathrm{MSE} = 1832.8, p < .001$, $F2(1,68) = 4.37, \mathrm{MSE} = 3580.1, p < .05$; type of word formation: $F1(1,29) = 18.2, \mathrm{MSE} = 3573.7, p < .001$ $F2(1,68) = 12.8, \mathrm{MSE} = 3573.7, p < .001$; interaction: $F1(1,29) = 3.4, \mathrm{MSE} = 2380.2, p > .05$, $F2(1,68) = 1.0, \mathrm{MSE} = 3573.7, p > .3$). In the error analysis, Family Size and Type of Word Formation emerged as reliable factors in the by-participant analysis and as marginally reliable in the by-item analysis (Family Size: $F1(1,29) = 16.0, p = 0.001, \mathrm{MSE} = 0.0053$; $F2(1,76) = 3.12, p = 0.081, \mathrm{MSE} = 0.0171$; Type of Word Formation: $F1(1,29) = 13.2, p = 0.001, \mathrm{MSE} = 0.0057$; $F2(1,76) = 3.07, p = 0.084, \mathrm{MSE} = 0.0171$). Their interaction was again absent ($F1, F2 < 1$). The main effect of Type of Word Formation is probably due wholly or in part to the difference between the average Surface Frequencies for inflected and derived words in the materials. The main effect of Family size in combination with the absence of a reliable interaction clearly shows that effects of Family Size can be equally strong for inflected and derived words.

These results lead us to expect to find a solid effect of Family Size as well for the derivational suffix -heid, which creates abstract nouns from adjectives (goedheid, 'goodness').

Experiment 6

Method

Participants. Sixteen students participated in the lexical decision experiment. Sixteen different students did the subjective frequency rating.

Materials. Forty derived words with the de-adjectival suffix -heid were selected from the CELEX database, of which twenty had a large Family Size (mean 37.0, median 26.5, range 14 – 132, full form Family Size 0.5), whereas the other twenty had a small Family Size (5.9, median 4.5, range 2 – 28, full form Family Size 0.1). The two sets were matched for Base Frequency (high Vf: 75; low Vf: 74), Surface Frequency (high Vf: 1.7; low Vf: 1,5), geometric mean bigram frequency (high Vf: 13.4; low Vf: 13.2), and word length in letters (high Vf: 8.5; low Vf: 8.9). The materials are listed in the Appendix. An additional 100 filler words consisted of monomorphemic and inflected nouns, adjectives and verbs, and derivations with the deverbal suffixes -sel and -ing. For each of the 140 words, a nonword was obtained by changing one to three letters, such that the phonotactics of Dutch were not violated.

Procedure. The procedure was identical to that of Experiment 1.

Results and Discussion

For the lexical decision task, the data of all participants were included in the analyses, as all performed with an overall error rate lower than 15%. Similarly, all items were included, as no item elicited error rates higher than 30%. The observations were used to calculate mean response latencies and error scores for the different test conditions. Together with the rating scores, they are listed in Table 5.

PLACE TABLE 5 APPROXIMATELY HERE

Surprisingly, neither a paired t-test for participants nor a standard two-sample t-test for items showed a significant difference in the response latencies in visual lexical decision between the two item groups ($t1(15) = 1.34, p > .1; t2(38) = 1.02, p > .1$), and the same holds for the error scores ($t1, t2 < 1$). The subjective frequency ratings also did not reveal a significant difference in the expected direction ($t1(15) = 4.15, p > .9; t2(38) = 1.18, p > .8$).

We therefore checked whether there is any evidence in post-hoc correlation and regression analyses for an effect of Family Size for words in -heid. As a point of comparison, we used the results of Experiment 1 for words in -te, for which we observed an effect of Family Size.

Interestingly, while inflected words in -te reveal a reliable correlation of Family Size with RT ($r = -0.39, t(37) = -2.6, p < .02$), the words in -heid do not show a fully significant correlation ($r = -0.22, t(38) = -1.4, p > .08$). Similarly, Family Size emerges as a significant component of response latencies in a stepwise multiple regression analysis of words in -te (with a coefficient of $-16.9, p < .05$) whereas Family Size is removed from the model for words in -heid, leaving only a significant effect of Surface Frequency. Finally, we note that there is no significant effect in the error data either ($t2(38) = -0.75, p > .2$). This suggests that Family Size does not, or only weakly, influence lexical processing of words in -heid.

Note that Experiments 1 and 6 have the same power in terms of number of participants and the number of words. However, the two experiments differ slightly with respect to the contrast in Family Size, $36/3 = 12$ for Experiment 1 and $37/6 = 6.2$ for Experiment 6. By removing the 6 words with the highest Family Size from the high condition of Experiment 1, the ratio between the two conditions becomes $20.6/3.3 = 6.2$ as well. Due to the reduction in power as a consequence of the removal of 30% of the items of the high Family Size condition, a t-test on the reduced data set of Experiment 1 does not support the 32 ms difference between the two conditions as reliable ($t2(31) = -1.38, p = 0.18$). However, Family Size still correlates reliably with RT ($r = -0.36, t(31) = -2.13, p = 0.02$), and Family Size together with Base Frequency remain as reliable factors in a stepwise linear regression analysis in which Surface Frequency is eliminated. Given that Family Size does not play a reliable role in the post-hoc regression and correlation analyses of Experiment 6, we therefore conclude that the different outcomes of Experiments 1 and 6 cannot be attributed to differences in the power of the two experiments.

Although we have not observed a statistically reliable effect of Family Size for -heid, either in our factorial experiment, in a regression analysis of Family Size and RT, or in the subjective frequency ratings, and although the power of the experiment does not appear to be at issue, the effect of Family Size is nevertheless in the predicted direction, as pointed out by our reviewers. To ascertain whether perhaps a more sophisticated count of the number of morphological family members might give rise to a reliable correlation of Family Size and RT, we first consider the role of semantic transparency of the family members, and then examine the role of the syntactic word categories of these family members.

Up till now, our count of family members is based on all formations listed in the CELEX lexical database. However, Schreuder and Baayen (1997) report for monomorphemic words that the correlation between Family Size and RT improves by removing semantically opaque family members from the count. We therefore examined the 755 family members of our target words in -heid with respect to their semantic transparency. Each word was judged independently by three raters, and only those words were accepted as semantically opaque that were judged to be opaque by all three. For instance, gemeenheid, 'meanness', has as one of its family members the word gemeente, 'municipality'. This family member was removed from the counts. In this way, we partioned the set of family members into a subset of 511 transparent family members and a subset of 244 opaque family members. For each of our 40 target words, we calculated two new counts, a count of its transparent family members, and a count of its opaque family members. We then inspected the correlation of RT with these new counts. For the opaque family count, $r = 0.04$ ($t(38) < 1$) and $rs = -0.08$ ($Z = -.49, p = 0.62$). For the transparent family count, however, $r = -0.32$ ($t(38) = -2.09, p = 0.043$) and $rs = -0.30$ ($Z = -1.89, p = 0.059$).

In order to ensure that the absence of a reliable correlation for the opaque family members is not due to the smaller number of opaque family we ran a randomization test in which the 511 'transparent' labels and the 244 'opaque' labels were randomly permuted across the 755 family members. We found that the observed correlation was lower than any of the correlations in 1000 permutation runs. Conversely, we observed that the empirical correlation for the transparent family members was higher than any correlation in the same set of 1000 permutation runs (see Baayen, Schreuder, Bertram, & Tweedie, 1999, for further details). This shows that the absence of a correlation for the opaque family members is not due to the reduced size of this set: random subsets of 244 family members give rise to much higher correlations. We conclude that there is a reliable correlation of RT and Family Size when we count the semantically transparent family members only.

Continuing with these transparent family members, we discovered that the syntactic word category (Noun, Verb, Adjective) of the family members also plays a role. We partitioned the family members into those to which -heid does not attach, the nouns and the verbs, on the one had, and those to which -heid does attach, the adjectives. The set of nouns and verbs contained 360 family members, and the set of adjectives contained 151 family members. We expected to find that the family members of the word category to which -heid attaches would show improved correlations compared to the nominal and verbal family members. Surprisingly, we observed the reverse: For the nouns and verbs, $r = -0.41$ ($t(38) = -2.79, p = 0.008$) and $rs = -0.39$ ($Z = -2.43, p = 0.015$), while for the adjectives, $r = -0.07$ ($t(38) = -0.46, p = 0.65$) and $rs = -0.05$ ($Z = -0.34, p = 0.74$). By excluding the adjectives, we obtain an improved correlation that is fully reliable in both the parametric and the nonparametric correlation analysis, and that is supported at the 5% significance level as well by a randomization test.

This result is surprising and counterintuitive. Why would a suffix that is subcategorized for attaching to adjectives not allow adjectives in the family to become active? Possibly, there are adjectives in the morphological families of our words in -heid to which -heid does not attach. We therefore inspected the individual adjectives in the families. Interestingly, there are indeed many adjectives to which -heid cannot be attached, such as color compounds (zeegroen, 'sea green') and intensified adjectives (steenkoud, 'stone cold', i.e., 'very cold'). In fact, none of the 3017 formations in -heid listed in our database in CELEX has a color compound as base word. There are 19 instances of -heid attaching to a intensified adjective in our database, but 18 of these occur once only in the underlying corpus of 42 million word tokens. This shows that -heid does not, or only very reluctantly, attach to color compounds and intensified adjectives. While the proportion of incompatible adjectives (19) among the total number of adjectives which are attested with -heid in our CELEX-derived database (3017) is low: $19/3017 = 0.0063$, the proportion of incompatible adjectives among the total number of pooled adjectival family members of our experimental target words is much higher: $75/151 = 0.497$ ($X^2_{(1)} = 1184.16, p < .001$).

This overrepresentation of incompatible adjectives in the families of our -heid materials suggests that the adjectives that are incompatible with the selection restrictions of -heid are not activated in the central lexicon. This hypothesis can be tested by studying the correlations of the counts of incompatible adjectives in the family with RT and likewise the correlation between the counts of compatible family members and RT. For the materials of the present experiment, this partitioning leads to counts that are too small to allow reliable correlations to be observed ($t < 1$ for both correlation analyses). However, when we combine the present materials with the materials of the two experiments with -heid formations reported in Bertram et al. (1999) (see also Baayen et al., 1999), we observe a reliable correlation of the family counts based on the compatible adjectives with RT ($r = -0.36, t(99) = -3.87, p = 0.0002$; $rs = -0.29, Z = -2.88, p = 0.004$), while the corresponding correlation for the incompatible adjectival family members is not statistically reliable ($r = -0.15, t(99) = -1.49, p = 0.14$; $rs = -0.11, Z = -1.14, p = 0.25$). In the corresponding randomization test, the empirical correlation for the incompatible adjectives was lower than any of the correlations observed for the same number of family members in 1000 permutation runs, while the correlation observed for the compatible adjectives was higher than observed for 97.5% of the correlations in the same 1000 permutation runs.

The finding that only adjectives compatible with -heid are activated in the mental lexicon shows that the effect of Family Size is modulated by the morphological context in which a base word occurs.

General Discussion

This paper addresses the question whether the Family Size of a complex word co-determines visual lexical decision times. To answer this question, we have studied four Dutch affixes separately, ranging from a prototypical inflectional suffix (the past tense suffix -te) via an affix that straddles the boundary of inflection and derivation (the comparative suffix -er) to the derivational suffixes -er ('-er') and -heid ('-ness'). For each of these affixes, we have indeed observed an effect of Family Size. For the inflectional suffixes -te and -er, a simple factorial contrast already was sufficient to obtain reliable effects. For the derivational suffix -er, a large regression design was necessary for the effect to emerge reliably in the reaction times. For the derivational suffix -heid, a factorial design did not reveal a reliable effect, neither in the error data nor in the reaction time data. Likewise, this suffix was the only one that did not reveal a reliable effect of Family Size in the off-line subjective frequency ratings. However, the effect in the response latencies was in the predicted direction, which led us to examine the counts of the family members of our target words in greater detail. By removing semantically opaque family members as well as adjectival family members with which -heid does not combine from the set of family members, new counts of the Family Size were obtained that now correlated reliably with the response latencies. It is surprising to find that such subtly defined subsets of morphologically complex family members that are not themselves present in the visual input apparently are co-activated in the mental lexicon.

These results have four theoretical implications. First, our results show that the effect of Family Size does not depend on the activation of the base word itself. Comparatives in -er reveal a solid Family Size effect while at the same time Bertram et al. (1999) report the absence of a Frequency effect for this suffix. Apparently, the type frequency effect of Family Size is independent of the token frequency effect of Base Frequency. Combined with the observation that the Family Size effect for our materials in -heid appears to be driven exclusively by the semantically transparent family members, we now have strong empirical support for the hypothesis of Schreuder and Baayen (1997) that the Family Size effect arises due to activation spreading between semantically transparent morphologically related words stored in the central lexicon.

Second, the observation that for our materials in -heid the Family Size effect appears to be driven by the semantically transparent family members shows that large numbers of semantically transparent and otherwise also completely regular complex forms must be stored in some way in the mental lexicon. This observation runs counter to the model advanced by Pinker (1991) and Clahsen et al. (1997) that only formally or semantically unpredictable formations are stored in the mental lexicon. This observation also implies that the hypothesis of strict decomposed storage in the central lexicon, with only morphemes receiving independent representation (Caramazza et al., 1988), cannot be true for the derived words and compounds in our experimental families. The mere existence of the Family Size effect as a type count effect shows that knowledge of the co-occurrence of two and often more than two morphemes in transparent complex words is part of what is stored in the mental lexicon.

Third, the present data show that the distinction between inflection and derivation as such is too coarse to predict whether an effect of Family Size will be observed: We observe the effect for both the inflectional and the derivational affixes in this study. There is one property specific to derivation, however, that we have found to play a role, namely, that derivational affixes may be subject to semantic selection restrictions. While inflectional affixes apply across the board without reference to the meaning of the base word, derivational suffixes may be restricted with respect to the semantics of the base word. Words in -heid do not attach to adjectives that specify a specific point or region on the scale they refer to, such as color compounds ('sea-green') and intensified adjectives ('stone-deaf'). Specific adjectives are overrepresented in the families of our target words in -heid, and we suspect that this overrepresentation underlies the observed absence of a reliable correlation of the raw counts of the adjectival family members with the response latencies. This result, which clearly requires replication studies,[4] suggests that the Family Size effect has a truly morphological component. The morphological context in which a word appears influences the way in which semantic activation spreads in the mental lexicon.

In the model developed by Schreuder and Baayen (1995), we can account for the absence of activation of incompatible family members by introducing mutually inhibitory semantic nodes, a SPECIFIC node and a NON-SPECIFIC node. The node of the suffix -heid is connected with the NON-SPECIFIC node, so that activation of -heid will lead to activation of the NON-SPECIFIC node, which in turn will lead to inhibition of the SPECIFIC node and all adjectives that are connected with this SPECIFIC node such as color compounds and intensified adjectives. Note that in production a reverse process may take place. Once the conceptualization process has led to the activation of a a color compound or an intensified adjective, the SPECIFIC node will be activated and will inhibit the NON-SPECIFIC node and hence the -heid node with which it is in turn connected. Thus, the same mechanism that in production ensures that incompatible adjectives are not combined with -heid is mirrored in comprehension by the absence of any co-activation of incompatible family members.

Fourth, the existence of a Family Size effect for complex words calls for re-examination of previously reported results on Base Frequency and Family Frequency. For instance, Experiment 2 of Taft's (1979) seminal study contrasts inflected words matched for Surface Frequency with respect to what we have called Base Frequency. Using the counts in the CELEX lexical database, his high Base Frequency condition turns out to be also a high Family Size condition (Base Frequency: 5913 versus 546 per 18 million, Family Size: 13.3 versus 3.3). Both log Base Frequency and log Family Size correlate reliably with the response latencies reported by Taft (Base Frequency: $r = -0.48, t(38) = -3.37, p = 0.0017$; Family Size: $r = -0.49, t(38) = -3.46, p = 0.0014$). As Family Size and Base Frequency are highly correlated (in Taft's data set, $r = 0.61, t(38) = 4.72, p < .0001$), it is not clear which of these factors is in fact measured in this experiment. In order to ascertain whether Base Frequency and Family Size play an independent role in English inflection, factorial designs as used in the present study, in which the covarying factor is controlled for, are necessary. The same caution is required with respect to the study of Bradley (1979), who reports Base Frequency effects for suffixed words in -er, -ment, and -ness, and with respect to the Base Frequency effects of Burani and Caramazza (1987). Our present results suggest that it is necessary to check whether these interesting findings are not due to a confound with Family Size in their materials.

Our results suggest that there are two kinds of frequency effects in human memory. On the one hand, there are token-frequency effects, which concern individual words by themselves. A high token frequency ensures that a given word is well-established in memory. On the other hand, the type frequency effect of Family Size concerns words, both simplex and complex, in morphological relation to other words in the mental lexicon. Interestingly, the functionality of morphology as a linguistic device is to transcend the arbitrary relation of form and meaning that characterizes simplex words. Most words (in the dictionary sense) in human languages are morphologically complex, using a relatively small number of simplex words and affixes to build large numbers of complex words with non-arbitrary mappings of form and meaning. The ubiquitous use of morphology results in large partially overlapping networks of morphologically related words. In fact, we suspect that the interconnectivity in the human mental lexicon is largely carried by these networks of morphological relations. From a cognitive point of view, the storage of complex words not as islands but as representations sharing aspects of form and meaning with many other representations (Schreuder & Baayen, 1995) is functional in that it leads to a considerable reduction of memory load. We interpret the Family Size effect as empirical evidence for such substantial interconnectivity in the mental lexicon along lines of morphological relatedness.

References

Anderson, S.R. (1992). A-morphous morphology, Cambridge: Cambridge University Press.

Baayen, R. H., Burani, C. and Schreuder, R. (1997). Effects of semantic markedness in the processing of regular nominal singulars and plurals in Italian. In G. E. Booij and J. v. Marle (Eds.), Yearbook of Morphology 1996 (pp. 13–34). Dordrecht: Kluwer Academic Publishers.

Baayen, R. H., Dijkstra, T. and Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual route model. Journal of Memory and Language 36, 94–117.

Baayen, R. H., Lieber, R. and Schreuder, R. (1997). The morphological complexity of simplex nouns. Linguistics 35, 861–877.

Baayen, R. H., Piepenbrock, R. and Gulikers, L. (1995). The CELEX lexical database (CD-ROM). Philadelphia, PA.: Linguistic Data Consortium, University of Pennsylvania.

Baayen, R.H., Schreuder, R., Bertram, R. and Tweedie, F.J. (1999). The semantic functions of -heid. Submitted.

Belsley, D.A., Kuh, E. and Welsch, R.E. (1980). Regression Diagnostics. Identifying Influential Data and sources of Collinearity. New York: Wiley.

Bertram, R., Schreuder, R. and Baayen, R. H. (1997). The balance of storage and computation in morphological processing: the role of Kind of Word Formation, Affixal Homonymy, and Productivity. Submitted.

Booij, G. E. (1993). Against split morphology. In G. E. Booij & J. v. Marle (Eds.), Yearbook of morphology 1993 (pp. 27–49). Dordrecht: Kluwer Academic Publishers.

Booij, G. E. (1998). The demarcation of inflection: a synoptical survey. Fabri, R., Ortmann, A. and Parodi, T. (Eds.) Models of Inflection (pp. 11–27). Tuebingen: Max Niemeyer Verlag.

Bradley, D. C. (1979). Lexical Representation of Derivational Relation. In: M. Aronoff and M.L. Kean (Eds.), Juncture. (pp. 37–55). Saratoga: Anma Libri.

Burani, C. and Caramazza, A. (1987). Representation and processing of derived words, Language and Cognitive Processes, 2, 217–227.

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), Language production (Vol. II): Development, writing and other language processes (pp. 257–294). London: Academic Press.

Bybee, J. L. (1985). Morphology: A study of the Relation between Meaning and Form. Amsterdam: Benjamins.

Caramazza, A., Laudanna, A. & Romani, C. (1988). Lexical access and inflectional morphology. Cognition, 28, 297–332.

Clahsen, H., Eisenbeiss, S., and Sonnenstuhl-Henning, I. (1997). Morphological structure and the processing of inflected words, Theoretical Linguistics 23, 201–249.

De Jong, N.H., Schreuder, R. and Baayen, R.H. (1999). Family size effects for uninflected and inflected verbs. Submitted.

Grainger, J. and Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model, Psychological Review 103, 518–565.

Grainger, J. and Segui, J. (1990). Neighborhood frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencies, Perception and Psychophysics 47, 191–198.

Hasher, L. and Zacks, R. T. (1984). Automatic Processing of Fundamental Information. The Case of Frequency of Occurrence, American Psychologist 39, 1372–1388.

Jescheniak, J.D. and Levelt, W.J.M. (1994). Word frequency effects in speech production: retrieval of syntactic information and of phonological form, Journal of Experimental Psychology: Learning, Memory and Cognition 20, 824–843.

Laudanna, A. and Burani, C. (1995). Distributional properties of derivational affixes: Implications for processing. In L. B. Feldman (Ed.), Morphological Aspects of Language Processing (pp. 345–364). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Lyons, J. (1968). Introduction to Theoretical Linguistics. Cambridge: Cambridge University Press.

Niemi, J. and Laine, M. and Tuominen, J. (1994) Cognitive morphology in Finnish: foundations of a new model, Language and Cognitive Processes 9, 423-446.

Oldfield, R. C. and Wingfield, A. (1965). Response Latencies in Naming Objects, Quarterly Journal of Experimental Psychology 17, 273–281.

Pinker, S. (1991). Rules of language. Science 153, 530–535.

Rubenstein, H. and Pollack, I. (1963) Word predictability and intelligibility, Journal of Verbal Learning and Verbal Behavior 2, 147–158.

Schreuder, R. & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), Morphological aspects of language processing (pp. 131–154). Hillsdale, New Jersey: Lawrence Erlbaum.

Schreuder, R. and Baayen, R. (1997). How complex simplex words can be, Journal of Memory and Language 36, 118–139.

Seidenberg, M. (1987). Sublexical structures in visual word recognition: Access units or orthographic redundancy. In M. Coltheart (Ed.), Attention and performance XII (pp. 245–263). Hove: Lawrence Erlbaum Associates.

Sereno, J. and Jongman, A. (1997). Processing of English Inflectional Morphology, Memory and Cognition 25, 425–437.

Taft, M. (1994). Interactive-activation as a Framework for Understanding Morphological Processing, Language and Cognitive Processes 9, 271–294.

Taft, M. (1979). Recognition of Affixed Words and the Word Frequency Effect, Memory and Cognition 7, 263–272.

Van Jaarsveld, H. J., Coolen, R. and Schreuder, R. (1994). The role of analogy in the interpretation of novel compounds, Journal of Psycholinguistic Research 23, 111–137.

Venables, W.N. and Ripley, B.D. (1995). Modern Applied Statistics with Splus, New York: Springer.

Whaley, C. P. (1978). Word–Nonword classification time, Journal of Verbal Language and Verbal Behavior 17, 143–154.

Author Notes

We wish to thank three anonymous reviewers, Joan Sereno, Matti Laine and Jukka Hyönä for their helpful comments and criticisms on earlier versions of this paper, Leonore Biegstraten for her help in running the experiments, and Dafydd Gibbon and James McQueen for stimulating discussions. The first author was financially supported by the Academy of Finland (grant #27774 to Matti Laine), and the study was also supported by the Dutch National Research Council NWO (PIONIER grant to the second author). Requests for reprints should be addressed to Raymond Bertram, Department of Psychology, University of Turku, FIN-20520 Turku, Finland. E-mail: Rayber@utu.fi.

Footnotes

[1] The one study that we are aware of that calls attention to a possible role of a type count of family members is Van Jaarsveld, Coolen and Schreuder (1994), who showed that a large Family Size of the constituents of novel nominal compounds made it more difficult to decide that these words were not already existing words. This is the mirror pattern of what we have observed for lexical decision on existing words: A large Family Size renders a string more word-like.

[2] In lexical statistics, N is often used to denote the number of tokens and V the number of types. We use Vf for the number of types in the morphological family.

[3] The Base Frequency is defined as the summed frequencies of the base word itself and its inflectional variants (for experimental evidence that the Base Frequency is a better predictor of response latencies in visual lexical decision and subjective frequency ratings than the Surface Frequency of the base itself, see Baayen, Dijkstra & Schreuder, 1997; Baayen, Burani & Schreuder, 1997; Schreuder & Baayen, 1997).

[4] A recent study reporting a facilitatory effect of a selection restriction on a subset of verbal family members using a controlled experimental design is De Jong, Schreuder, and Baayen (1999).

Appendix

Materials used in the experiments Each word is followed by mean RT in ms., mean error score (in %) and mean subjective frequency rating where available. Words marked with a † were not included in the analyses due to their high error rate

Words Used in Experiment 1

Verbs with a Large Family Size: bokste (boxed): 592 7.1 4.0; braakte (vomited): 652 0 3.6; haakte (hitched): 649 6.2 3.3; kapte (chopped): 660 0 3.8; klapte (clapped): 635 0 4.8; klopte (knocked): 607 0 5.9; knoopte (knotted): 714 7.1 5.0; kruiste (crossed): 737 6.2 3.8; lichtte (lighted): 548 6.2 4.1; lustte (liked): 690 0 5.8; pompte (pumped): 674 0 3.4; raapte (picked): 662 7.1 4.3; rookte (smoked): 596 0 6.4; schepte (created): 665 0 4.1; schraapte (scraped): 708 0 3.8; spitste (pricked up): 736 0 3.6; startte (started): 625 6.2 5.9; stootte (nudged): 702 7.1 5.1; stroopte (poached): 739 7.1 3.2; twistte (quarreled) 716 0 2.8

Verbs with a Small Family Size: wreekte (revenged)† : 812 31.2 2.8; bukte (stooped): 655 7.1 4.7;

dempte (dimmed): 797 12.5 2.8; grinnikte (chuckled): 659 0 4.4; hurkte (squatted): 689 6.2 4.1; juichte (cheered): 694 0 4.8; ketste (turned down): 863 12.5 2.9; krenkte (offended): 725 0 2.9; krijste (shrieked): 733 6.2 4.6; mikte (aimed): 658 0 4.1; repte (mentioned): 735 14.3 2.7; schrapte (struck out): 732 14.3 2.9; siste (hissed): 713 14.3 2.3; slikte (swallowed): 588 6.2 5.4; snikte (sobbed): 575 7.1 4.4; snurkte (snored): 681 7.1 4.8; strekte (stretched): 648 6.2 4.6; tartte (defied): 880 18.8 2.9; trachtte (attempted): 676 0 4.6 zwiepte (swished): 646 14.3 2.3

Words Used in Experiment 2

Comparatives with a Large Family Size: blauwer (bluer): 648 6.2 3.4; boller (chubbier): 665 0 4.1; fijner (finer): 532 0 6.2; groener (greener): 555 0 3.9; korter (shorter): 559 0 6.6; losser (looser): 582 0 5.3; mobieler (mobiler): 573 0 4.3; netter (more decent): 612 12.5 5.6; platter (flatter): 537 6.2 5.2; rijker (richer): 521 0 5.8; rotter (more rotten): 647 6.2 3.8 schaarser (rarer): 619 0 3.2; spitser (sharper): 719 6.2 3.9; veiliger (safer): 543 0 5.9; vetter (greasier): 599 0 5.2; vlakker (flatter): 582 12.5 4.0; voller (fuller): 594 0 5.8; vuiler (dirtier): 543 6.2 5.4; zouter (salter): 592 6.2 4.3; zwakker (weaker): 568 0 5.7

Comparatives with a Small Family Size: banger (more afraid): 567 6.2 5.6; feller (fiercer): 678 0 5.3; flinker (firmer): 716 0 4.3; gammeler (seedier): 838 12.5 3.2; gauwer (sooner): 976 18.8 3.4; geringer (smaller): 620 0 3.8; ijdeler (vainer): 584 0 3.3; jaloerser (more jealous): 722 0 3.7; juister (righter): 590 0 4.8; lauwer (more tepid): 704 18.8 2.8; leuker (nicer): 511 0 6.6; milder (milder): 612 6.2 4.1; moderner (more modern): 738 0 5.6; mooier (more beautiful): 549 0 6.8; norser (gruffer): 721 6.2 3.1; schriller (shriller): 735 25 1.9; schuiner (more oblique): 633 0 4.8; simpeler (simpler): 614 0 5.1; steiler (steeper): 577 0 5.0; trotser (prouder): 692 0 4.7

Words Used in Experiment 3

Nouns in -er with a Large Family Size: schender (violator)† : 897 31.2 3.3; wringer (wringer)†: 760 31.2 3.3; klever (adheror): 660 25.0 2.8; schuiler (hider): 706 25.0 2.3; smelter (melter): 695 25.0 1.7; draver (trotter): 638 18.8 4.1; strever (striver): 581 18.8 4.6; buiger (bower): 755 6.2 1.8; gieter (watering-can): 548 6.2 6.3; slijter (licensed victualler): 615 6.2 6.6; zwemmer (swimmer): 586 6.2 6.6; blusser (extinguisher): 586 0 5.3; kruiper (creeper): 581 0 3.4; lader (loader): 680 0 4.4; rijder (rider): 555 0 5.8; roeier (rower): 567 0 5.7; slijper (polisher): 589 0 4.8; sluiper (sneaker): 591 0 4.1; strooier (strewer): 557 0 4.1; vriezer (freezer): 630 0 6.6

Nouns in -er with a Small Family Size: tarter (defier)† : 903 75 1.4; schreier (weeper)†: 858 63 1.6; bader (bather)† : 946 50 2.5; pocher (boaster)†: 754 50 3.2; delver (digger)† : 784 38 3.6; peinzer (meditator)†: 688 38 3.7; glimmer (shiner)† : 742 31 2.0; janker (yelper)†: 752 31 3.0; tobber (worrier)† : 704 31 4.8; peller (peeler): 784 25 2.9; bruller (roarer): 608 19 3.1; glijder (slider): 569 19 2.9; demper (damper): 740 13 3.9; dwinger (forcer): 592 13 2.0; bieder (bidder): 602 6 4.9; ruiker (bouquet): 511 6 4.9; sisser (hisser): 747 6 4.8; smijter (dasher): 622 6 2.3; streler (caresser): 672 6 3.0; wreker (revenger): 618 0 3.9

Words Used in Experiment 4

bader (bather)†: 706 47.1; bidder (prayer): 680 11.8; bieder (bidder): 607 5.9; binder (binder): 683 5.9; blusser (extinguisher): 613 0.0; breker (breaker): 564 5.9; brenger (bringer): 602 23.5; bruller (howler): 633 17.6; buiger (bender): 599 11.8; delver (digger): 701 11.8; domper (extinguisher): 689 17.6; drager (carrier): 584 0.0; draler (dawdler): 684 29.4; drammer (nag): 625 23.5; draver (trotter): 597 5.9; drenzer (whiner)†: 575 88.2; drijver (driver): 613 0.0; drinker (drinker): 547 0.0; duider (pointer)†: 792 41.2; dweper (devotee): 565 23.5; dwinger (forcer): 643 0.0; eter (eater): 601 0.0; gapper (thief)†: 941 52.9; gieter (watering can): 595 0.0; glijder (glider): 603 0.0; griener (wailer)†: 827 35.3; heerser (ruler): 560 0.0; helper (helper): 515 5.9; hijger (heavy breather): 572 0.0; houder (holder): 592 0.0; huiler (weeper): 572 5.9; jager (hunter): 537 0.0; janker (sniveler): 608 17.6; keffer (barker): 725 29.4; kenner (connoisseur): 567 0.0; kermer (sniveler): 805 23.5; klager (complainer): 547 0.0; kleumer (shivery person)†: 811 47.1; knager (gnawer): 592 5.9; kneder (molder): 661 17.6; kniezer (moaner): 745 17.6; kruier (porter): 693 11.8; kruiper (crawler): 575 5.9; lader (loader): 655 23.5; leider (leader): 523 0.0; lezer (reader): 521 0.0; ligger (girder): 625 5.9; lijder (sufferer): 613 0.0; nemer (buyer)†: 649 41.2; peinzer (meditator): 625 5.9; pletter (flatter): 628 5.9; pocher (boaster)†: 737 58.8; prediker (preacher): 628 0.0; pruiler (pouter): 790 23.5; ravotter (romper): 730 17.6; redder (rescuer): 534 0.0; reder (shipowner): 747 5.9; richter (judge): 634 17.6; roeier (rower): 585 5.9; ronker (snorer)†: 916 82.4; ruiker (bouquet): 581 5.9; schelder (curser)†: 711 41.2; schender (breaker)†: 798 41.2; schreier (weeper)†: 865 58.8; schrijver (writer): 528 0.0; schuiler (hider): 608 23.5; schutter (archer): 599 0.0; sisser (hisser): 680 11.8; slijter (licensed victualler): 546 0.0; sluiter (shutter): 658 23.5; smijter (thrower): 568 5.9; snijder (cutter): 593 5.9; snoever (swaggerer)†: 660 47.1; speler (player): 542 0.0; spieker (copier): 633 5.9; spreker (speaker): 558 0.0; spuwer (spitter): 591 17.6; stijger (riser): 629 5.9; stinker (stinker): 628 5.9; stoker (fireman): 611 0.0; stouwer (lumper)†: 690 47.1; streler (caresser): 550 5.9; strever (striver): 610 17.6; strooier (scatterer): 596 0.0; tarter (provoker)†: 737 64.7; teler (cultivator): 609 0.0; temmer (tamer): 694 29.4; tobber (worrier): 618 11.8; trainer (trainer): 533 5.9; trimmer (jogger): 740 29.4; turner (turner): 612 11.8; vechter (fighter): 554 0.0; veinzer (dissembler): 711 29.4; vinder (finder): 718 0.0; vitter (faultfinder): 716 23.5; vleier (flatterer): 585 5.9; voeler (feeler): 694 5.9; volger (follower): 567 5.9; vreter (glutton): 578 5.9; vriezer (freezer): 580 0.0; vuller (filler): 563 11.8; waaier (fan): 609 0.0; wekker (alarm clock): 617 0.0; werper (thrower): 581 0.0; wever (weaver): 632 0.0; wieder (weeder): 694 29.4; winner (winner): 645 11.8; wreker (avenger): 594 0.0; wringer (wringer): 692 29.4; wroeter (burrower): 767 17.6; wurger (strangler): 599 11.8; zender (sender): 572 0.0; ziener (prophet): 571 11.8; zoemer (buzzer): 630 0.0; zuiger (piston): 541 0.0; zwalker (wanderer): 758 29.4; zwemmer (swimmer): 554 0.0; zwerver (drifter): 527 0.0; zwetser (boaster): 718 17.6; zwijger (silent person): 547 0.0; zwoeger (plodder): 596 0.0;

Words Used in Experiment 5

Inflected verbs with a Large Family Size: knoopte (tied up): 560 0.0; ontglip (slip away): 718 14.3; verstoot (cast out): 692 35.7; kapte (do one's hair): 677 25.0; bokste (boxed): 592 14.3; braakte (vomited): 605 0.0; doopte (baptised): 595 7.1; klapte (clapped): 562 0.0; klopte (knocked): 517 0.0; kruiste (cruised): 636 0.0; lichtte (lifted): 611 0.0; lustte (liked): 607 0.0; piepte (whined): 550 0.0; pompte (pumped): 656 0.0; rookte (smoked): 480 0.0; schepte (created): 566 0.0; schraapte (scraped): 613 7.1; spitste (pricked): 652 0.0; startte (started): 578 0.0; stootte (jolt): 601 0.0; stroopte (poached): 568 0.0; twistte (quarrel): 588 0.0;

Inflected verbs with a Small Family Size: bonkte (knocked): 556 14.3; bukte (stooped): 604 0.0; dempte (muffled): 708 7.1; glipte (slipped): 585 0.0; grinnikte (chuckled): 616 0.0; jankte (whined): 569 0.0; juichte (cheered): 557 0.0; knarste (crunched): 647 0.0; krenkte (offend): 610 7.1; laakte (reprehended): 813 62.5; mikte (aimed): 584 7.1; schiftte (sifted): 946 37.5; siste (sizzled): 532 0.0; slikte (swallowed): 533 0.0; snikte (gasped): 547 0.0; snurkte (snored): 598 0.0; strekte (stretched): 565 7.1; trachtte (attempted): 652 12.5; wreekte (avenged): 632 0.0; zwiepte (walloped): 736 12.5;

Derived words with a Large Family Size: betwist (dispute): 588 7.1; bokser (boxer): 562 0.0; braaksel (vomit): 638 0.0; doper (baptist): 729 50.0; geklap (clapping): 652 0.0; kapsel (haircut): 572 7.1; klopper (knocker): 612 14.3; knooppunt (junction): 564 0.0; kruiser (cruiser): 731 21.4; lichting (lifting): 597 0.0; lustig (cheerful): 578 0.0; pieper (whiner): 557 0.0; pompwater (pump water): 639 0.0; rokerij (smokehouse): 604 6.2; schepsel (creature): 584 0.0; schraapsel (scrapings): 849 25.0; starter (starter): 626 0.0; stroperij (poaching): 677 0.0; spitsuur (rush hour): 577 0.0;

Derived words with a Small Family Size: betracht (practise): 624 7.1; demping (muffling): 809 12.5; gebonk (thumping): 731 12.5; gegrinnik (chuckling): 707 0.0; gejuich (jubilation): 562 0.0; geknars (gnashing): 668 7.1; jankerig (whiny): 637 7.1; krenking (offence): 844 50.0; laakbaar (reprehensible): 704 14.3; mikpunt (target): 562 0.0; schifting (sifting): 737 7.1; sisser (squib): 674 12.5; snurker (snorer): 647 14.3; strekking (stretching): 614 0.0; verslik (choke): 678 7.1; wreker (avenger): 638 6.2; zwieper (walloper): 764 42.9; doodssnik (last gasp): 967 37.5; gebukt (weighed down): 627 7.1;

Words Used in Experiment 6

Nouns in -heid with a Large Family Size: blindheid (blindness): 651 6.2 5.3; domheid (stupidity): 611 0 5.6; droogheid (dryness): 612 0 3.9; duisterheid (darkness): 613 0 4.1; dwarsheid (crossness): 620 6.2 2.6; fijnheid (fineness): 628 0 2.8; gemeenheid (meanness): 737 6.2 3.2; hoogheid (highness): 567 6.2 4.8; matheid (weariness): 636 6.2 2.8; naaktheid (nudeness): 621 6.2 3.9; platheid (flatness): 655 0 2.1; rijpheid (ripeness): 643 0 4.9; rotheid (rottenness): 626 6.2 1.7; spitsheid (sharpness): 655 0 2.0; wijdheid (wideness): 775 6.2 3.1; wildheid (wildness): 736 0 2.2; zoutheid (saltness): 639 0 2.1; zuiverheid (cleanness): 583 0 5.6; zuurheid (sourness): 671 0 2.3; zwartheid (blackness): 649 0 2.1

Nouns in -heid with a Small Family Size: bangheid (timidity): 674 0 3.7; bitterheid (bitterness): 630 6.2 4.7; blankheid (whiteness): 640 0 2.0; felheid (fierceness): 678 6.2 4.8; flinkheid (boldness): 658 0 3.5; forsheid (robustness): 743 0 2.9; fraaiheid (beauty): 663 12.5 3.4; geringheid (smallness): 686 6.2 2.8; grijsheid (greyness): 646 0 2.6; juistheid (correctness): 639 12.5 6.3; leukheid (niceness): 735 0 2.3; luiheid (lazyness): 625 12.5 6.1; magerheid (thinness): 679 0 2.4; puurheid (purity): 657 0 3.4; saaiheid (dullness): 645 0 4.8; serieusheid (seriousness): 792 6.2 3.4; simpelheid (simplicity): 629 0 3.9; somberheid (gloomyness): 609 0 5.7; strakheid (stiffness): 632 6.2 2.6; traagheid (slowness): 589 0 5.8;

Table 1

Mean Response Latencies with Standard Deviation (SD),

Error Percentages and Subjective Frequency Ratings

with SD (by items) for Verbal Inflections with the Past Tense Suffix -te

with a Large and a Small Family Size (Experiment 1)

	Family Size	RT	SD	Error	Rating	SD
Large	36.0	665	54	3.0	4.3	1.0
Small	3.3	702	79	8.9	3.8	1.0

Table 2

Mean Response Latencies with SD, Error Percentages and Subjective Frequency Ratings with SD (by items) for Inflected Words with the Comparative Suffix -er with a Large and a Small Family Size (Experiment 2)

	Family Size	RT	SD	Error	Rating	SD
Large	52.2	589	51	3.1	4.9	1.0
Small	3.5	669	109	4.7	4.4	1.3

Table 3

Mean Response Latencies with SD, Error Percentages and Subjective Frequency Ratings with SD (by items) for Derived Words with the Deverbal Suffix -er with a Large and a Small Family Size (Experiment 3).

	Family Size	RT	SD	Error	Rating	SD
Large	23.1	618	59	10.0	4.4	1.6
Small	5.2	642	84	26.6	3.2	1.1

Table 4

Mean Response Latencies with SD and Error Percentages (by items) for Inflected and Derived Words with a Large and a Small Family Size (Experiment 5).

	Inflected			Derived		
	RT	SD	Error	RT	SD	Error
Large	587	43	1.5	624	73	4.5
Small	603	57	3.8	668	64	7.6

Table 5

Mean Response Latencies with SD, Error Percentages and Subjective Frequency Ratings with SD (by items) for Derived Words with the De-adjectival Suffix -heid with a Large and a Small Family Size (Experiment 6).

	Family Size	RT	SD	Error	Rating	SD
Large	37.0	646	51	2.5	3.4	1.3
Small	5.9	662	48	3.4	3.9	1.4