Running head: THE COMPLEXITY OF SIMPLEX WORDS

How complex simplex words can be

Robert Schreuder
Interfaculty Research Unit
for Language and Speech,
University of Nijmegen

R. Harald Baayen
Max Planck Institute
for Psycholinguistics,
Nijmegen

How complex simplex words can be.

Send proofs to:
R. Schreuder
Interfaculty Research Unit for Language and Speech
Wundtlaan 1
6525 XD Nijmegen
The Netherlands

Abstract

A series of experiments investigated components of the word frequency effect in visual lexical decision, progressive demasking, and subjective frequency ratings. For simplex, i.e., monomorphemic, nouns in Dutch, we studied the effect of the frequency of the monomorphemic noun itself as well as the effect of the frequencies of morphologically related forms on the processing of these monomorphemic nouns. The experiments show that the frequency of the (unseen) plural forms affects the experimental measures. Nouns with high-frequency plurals are responded to more quickly in visual lexical decision, and they receive higher subjective frequency ratings. However, the summed frequencies of the formations in the morphological family of a given noun (the compounds and derived words in which that noun appears as a constituent) did not affect the experimental measures. Surprisingly, the size of the morphological family, i.e., the number of different words in the family, emerged as a substantial factor. A monomorphemic noun with a large family size elicits higher subjective frequency ratings and shorter response latencies in visual lexical decision than a monomorphemic noun with a small family size. The effect of family size disappears in progressive demasking, a task which taps into the earlier stages of form identification. This suggests that the effect of family size arises at more central, post-identification stages of lexical processing.

One of the most robust findings in psycholinguistics is the word frequency effect (e.g., Gardner, Rothkopf, Lapan, & Lafferty, 1987; Gernsbacher, 1984; Gordon, 1983; Grosjean, 1980; Hasher & Zacks, 1984, Jescheniak & Levelt, 1994; Rubenstein & Pollack, 1963; Scarborough, Cortese, & Scarborough, 1977; Shapiro, 1969; Whaley, 1978). A high-frequency word such as *car* is recognized more quickly than a low-frequency word such as *doe*. For morphologically complex words, results have been reported that suggest that processing times are determined not only by the frequency of the complex word itself, but also by the frequencies of its constituents as these appear by themselves and in other words (Bradley, 1979; Burani, Salmaso, & Caramazza, 1984, Burani & Caramazza, 1987; Cole, Beauvillain, & Segui, 1989; Laudanna & Burani, 1985; Taft & Forster, 1976; Taft, 1979).

Taft (1979, Experiment 2) reported that reaction times in visual lexical decision to simplex, i.e., monomorphemic, words in English are codetermined by the frequencies of the inflectional variants of these words. Baayen, Dijkstra, and Schreuder (1997) have observed similar results for Dutch. Their experiments strongly suggest that the recognition time in lexical decision for singular nouns in Dutch is determined by the summed frequency of use of both their singular and their plural forms. Similar results for Italian nouns have been reported by Baayen, Burani, and Schreuder (1996). This pattern of findings indicates that, surprisingly, the speed with which monomorphemic nouns are recognized is determined not only by their own frequencies of use, but also by the frequencies of use of other, morphologically related, words such as plural inflections. Potentially, even their occurrences in derived words and compounds might codetermine the recognition latency of a monomorphemic noun.

In this paper, we investigate these potential components of the word frequency effect using visual lexical decision, progressive demasking, and subjective frequency ratings. We will show that in Dutch the frequency of the monomorphemic word itself, its so-called surface frequency, as well as the frequency of its inflectional variant, the plural, both codetermine reaction times and ratings. In addition, we will study how derived words and compounds affect the ratings and reaction times of their monomorphemic nominal constituents in Dutch. Surprisingly, our experiments show that the cumulated frequencies of derived words and compounds containing a given monomorphemic noun as a constituent do not affect lexical processing of the base noun, but that it is the number of different derived words and compounds that is crucially involved. (In Dutch, derivation proceeds along the same lines as in English. The same holds for compounding, be it that in Dutch all compounds are written as single words without intervening spaces, while in English only high-frequency, well-established compounds are written as single words.)

Thus, the word frequency effect, one of the most robust findings in psycholinguistics, is a composite effect in two respects. First, the word frequency effect for monomorphemic nouns is determined by the frequencies of occurrence of these monomorphemic nouns on the one hand, but on the other hand also by the frequencies of morphologically related complex words, words that are not themselves present in the visual input. Second, the word frequency effect is composite in nature in the sense that it has both a token and a type component. While a token count lies at the heart of the well-known word frequency effect, the role of a type count for the processing of monomorphemic words is a new finding in visual word recognition.

We will discuss the consequences of our findings for practical issues such as matching for word frequency in psycholinguistic experiments, and also for theoretical issues such as the nature of the word frequency effect and the organization of the mental lexicon, especially with respect to the processing of both morphologically simplex and morphologically complex words.

In this paper we present six experiments, all of which concern the processing of monomorphemic Dutch nouns. Experiment 1 considers the effect of the frequency of the plural form on the processing of the singular form. Experiment 2 broadens the scope by examining the potential effect of the token frequencies of derived words and compounds containing a given simplex noun as a constituent. Experiments 3 and 4 disentangle the effect of type versus token frequencies of these compounds and derived words. Experiment 5 focuses on the effect of the frequency of the singular form itself, and Experiment 6 studies the locus of the type frequency effect uncovered in Experiment 3.

Experiment 1

Experiment 1 is a partial replication study of Experiment 2 in Baayen, Dijkstra, and Schreuder (1997). These authors studied response latencies to singular nouns and their plurals in Dutch. They manipulated a number of factors, among them the surface frequencies of the singular and plural forms. In their Experiment 1, they kept the summed frequencies of these singulars and their corresponding plurals constant. Despite substantial differences in surface frequency, the singular nouns with a low surface frequency and the singular nouns with a high surface frequency were processed equally fast. Baayen, Dijkstra, and Schreuder (1997) argued that it is the summed frequency of the singular and plural inflections of a given stem, the so-called stem frequency, that crucially determines response latencies. In their Experiment 2, these authors show that sets of monomorphemic nouns which are matched for the surface frequency of the singular form but which differ with respect to the frequency of their plural form are not processed equally fast. The singulars with high-frequency plurals reveal shorter response latencies than the singulars with low-frequency plurals. This finding has serious consequences both for practical matters such as what kind of frequency counts are optimal for matching purposes, as well as for modeling the architecture of the mental lexicon. At the outset of the present study, we therefore first report a replication of this important result using an entirely different set of word materials, in which we focus on the processing of singular nouns only, keeping their surface frequencies constant, but varying the frequencies of their plural forms.

Method

Participants. Twenty-nine participants, mostly undergraduates at Nijmegen University, were paid to take part in the reaction time experiment, and twenty-nine different participants were paid to perform the subjective frequency rating experiment. All were native speakers of Dutch.

Materials. We selected our word materials from the medium surface frequency range of the singular form of approximately 100 to 600 occurrences per 42 million, using the CELEX lexical database (Baayen, Piepenbrock, & Van Rijn, 1993). Seventy nouns were selected. Half of these nouns had a plural form with a high surface frequency (mean 910 per 42 million), the other half had a plural form with a low surface frequency (mean 14 per 42 million). These two sets were matched for the frequency of the singular form, the number of syllables, word length in letters, and geometric mean bigram frequency. The summed frequency of singular and plural was 1218 per 42 million and 323 per 42 million respectively. The materials are listed in the Appendix. In addition, 10 monomorphemic practice words were selected from the same frequency range as the target words. Eighty phonotactically legal nonwords were constructed by changing one or two letters in monomorphemic words of the same frequency range.

Procedure. For the reaction-time experiment, participants were tested in groups of three in noise-proof experimental booths. They received standard lexical decision instructions. Each trial consisted of the presentation of a fixation mark (asterisk) in the middle of the screen during 500 ms, followed after 50 ms by the stimulus centered at the same position. Stimuli were presented on Nec Multisync color monitors in white upper-case 36 points Helvetica letters on a dark background. Stimuli remained on the screen for 1500 ms. Time-out occurred 2000 ms after stimulus onset. The total duration of the experiment was approximately ten minutes.

For the subjective frequency rating, participants were asked to indicate on a seven-point scale how often they thought a word is used in Dutch. The frequency range in our materials was relatively small. Therefore, we explicitly told our participants that most speakers of Dutch know the words in our list quite well, but that nevertheless there are differences in their frequency of use. As an example, we called attention to the words elbow and bus. Both words are well known, but elbow is a word that we generally do not use on a daily basis, while a word like bus is probably used somewhat more often in Dutch.

Results and Discussion

The results were fully in line with our predictions. Singular nouns with a high plural frequency received higher subjective frequency ratings and were responded to faster than equally frequent singulars with low frequency plurals. Table 1 shows that they were processed some 41 ms more quickly.

INSERT TABLE 1 APPROXIMATELY HERE

This difference in response latencies is highly reliable ($F1(1, 28) = 49.7, p < .001$; $F2(1, 68) = 14.4, p < .001$). Our participants performed the task without difficulties and with a high accuracy. Not surprisingly, the error analyses revealed no significant effects ($F1(1, 28) = 2.0, p > .10$; $F2 < 1$).

The difference in subjective frequency rating was also highly reliable. After removing the observations of one participant who failed to give a response to a large number of items, subject and item means were calculated. Analyses of variance by participants and by items revealed highly significant results ($F1(1, 27) = 306.8, p < .001$; $F2(1, 68) = 16.5, p < .001$).

Experiment 1 replicates the results reported in Baayen, Dijkstra, and Schreuder (1997), Experiment 2, in that monomorphemic noun singulars matched for surface frequency reveal substantially different processing times as a function of the frequency of the plural form, using different word materials. These results receive further confirmation by the subjective frequency rating, which is also highly sensitive to the frequency of the unseen plural form. Moreover, experiments in which the summed frequency of the singular and plural form is kept constant, and in which the frequency of the singular and plural forms is varied reveal that the singular forms are processed equally fast despite a substantial difference in surface frequency (see Baayen et al., 1996, for Italian, and Baayen, Dijkstra, & Schreuder, 1997, for Dutch). Considered jointly, we conclude that it is the summed frequency of the singular and plural forms that determines response latencies in visual lexical decision.

However, in addition to the plural form of a given monomorphemic noun, there are also the derived words and compounds in which that noun appears that potentially affect reaction times and subjective frequency estimates. At this point, it is useful to introduce some terminology.

We will use the term surface frequency to denote the frequency of use of a particular form. Thus, the surface frequency of the singular form table equals 3645 per 18 million, the surface frequency of the plural form tables is 563 per 18 million. We will also refer to these two frequency counts as the singular and plural frequencies. When we add the singular and plural frequency counts, we obtain what we will call the stem frequency, 4208 for table. The stem frequency of a word (in the sense of a dictionary entry) is the frequency of that word cumulating over all its inflectional variants. We will use the term morphological family to denote the set of words derived from a given stem by means of either compounding (tablespoon, timetable) or derivation (tablet, tabular). We will refer to the number of different words in the morphological family (excluding from the count the base word itself) as the morphological family size, and to the summed token frequencies of these words (now excluding the stem frequency of the base) as the cumulative family frequency. Note that the singular frequency, the plural frequency, and the cumulative family frequency are disjunct counts that jointly cover all word tokens in which the base occurs. In the next experiment, we explore the possible role of cumulative family frequency for monomorphemic nouns matched for stem frequency.

Experiment 2

Method

Participants. Twenty-nine participants, mostly undergraduates at Nijmegen University, were paid to participate in the reaction-time experiment. Thirty-one different participants performed the corresponding subjective frequency rating experiment. All were native speakers of Dutch. None of them had participated in Experiment 1.

Materials. We selected our word materials from a surface frequency range of approximately 10 to 1100 occurrences per 42 million, using the CELEX lexical database. Thirty-two nouns with a high cumulative family frequency (2680 per 42 million) were selected. In addition, thirty-two nouns with a very low cumulative family frequency (20 per 42 million) were obtained. These two sets were matched for the frequency of the singular, the frequency of the plural, the number of syllables, length in letters, and geometric mean bigram frequency. In addition, 10 monomorphemic practice words were selected from the same frequency range as the target words. Seventy-four phonotactically legal nonwords were constructed by changing one or two letters in monomorphemic words of the same frequency range.

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

Results and Discussion

For the reaction-time experiment, the data of three participants with an overall error score exceeding 10% were removed from the data set. As predicted, the singulars with a high cumulative family frequency were responded faster (45 ms) than the singulars with the same stem frequency but with a low cumulative family frequency, as shown in Table 2. These differences are highly significant, both for RT ($F1(1, 25) = 57.6, p < .001$; $F2(1, 62) = 9.9, p < .01$) and for error percentage ($F1(1, 25) = 40.5, p < .001$; $F2(1, 62) = 7.5, p < .01$).

INSERT TABLE 2 APPROXIMATELY HERE

For the subjective frequency rating, we also obtained a highly significant difference ($F1(1, 30) = 219.2, p < .001$; $F2(1, 62) = 8.2, p < .01$): the words with a high cumulative family frequency received a higher mean rating than the words for which this frequency is low (see Table 2).

These results suggest that indeed the cumulative family frequency of a monomorphemic noun codetermines its recognition latency. But do the cumulative family frequency and the stem frequency have equal weight? With respect to the weight of the inflectional variants, Baayen, Dijkstra, and Schreuder (1997) show that it is the summed frequency of both the singular and the plural inflectional variants that give optimal predictions. This holds not only for Dutch, where the singular form is monomorphemic, but also for Italian, where the singular form, just as the plural form, has an inflectional ending (Baayen et al., 1996). Thus it appears that the occurrences of inflectional variants of a noun have the same frequency weight as the occurrences of the base form itself. For the frequencies of morphologically related words, derived words and compounds, this does not appear to be true. In Experiment 1, a mean frequency of 14 for the low and 910 for the high plural frequency condition resulted in a 41 ms difference. But a much larger difference in mean cumulative family frequency, 20 for the low and 2680 for the high condition, resulted in a difference in reaction time of the same order of magnitude: 45 ms. This suggests that the importance of the cumulative family frequency is less than that of the inflectional frequencies.

To obtain some insight into the relative contribution of the cumulative family frequency, we carried out a series of post-hoc correlation analyses. Since Experiment 1 shows that the stem frequency is a better predictor of response latencies than the singular frequency, we first examined the correlation between stem frequency and reaction times for the high and low cumulative family frequency conditions separately. Here, and in all correlational analyses to follow, we have examined log frequency rather than absolute frequency, as absolute frequency is a non-linear predictor of response latencies, while log frequency is more linearly related to RT. (More precisely, we have used log (frequency+1), in order to be able to include 0 counts in our analyses.) As expected, the stem frequency revealed significant correlations with RT in the expected direction for both conditions (high, frequency range 14–1077: $r = -0.54, t(30) = -3.51, p < .001$; low, frequency range 18–1087: $r = -0.61, t(30) = -4.19, p < .001$). Turning to the cumulative family frequency, we found a surprising absence of a significant correlation with RT in the high condition: for a frequency range of 261–37420, $r = -0.26, t(30) = -0.25, p < .1$. (In this study, all t-tests concerning correlations between frequency measures and performance measures are one-tailed tests, as higher frequencies lead to faster response times and higher subjective frequency ratings.) Even more surprising, the low condition, with a much smaller frequency range of 0–107, yielded a highly significant correlation: $r = -0.46, t(30) = -2.87, p < .01$.

These post-hoc analyses pose a serious problem. How is it possible that the high cumulative family frequency condition does not reveal a significant effect on reaction time, even though the frequency range is much larger than that in the corresponding low condition? This contradictory pattern of results strongly suggests that the cumulative family frequency is confounded with another variable that is correlated with cumulative family frequency, but that has not been controlled for in the experiment. A possible variable that suggests itself is the size of the morphological family, the number of "morphological descendents" of a given stem.

Counts based on the CELEX lexical database show that family size and cumulative family frequency are indeed highly correlated in Dutch ($r = .78, t(22918) = 191.37, p < .001$, or, using a nonparametric correlation test (Spearman), $rs = .93, t(22918) = 140.97$, $p < .001$).

PLACE FIGURE 1 APPROXIMATELY HERE

Figure 1 displays the relation between family size V_f and cumulative family frequency N_f for Dutch morphemes (including both stems and affixes) using a logarithmic scale. "Hermit" morphemes, morphemes without any morphological descendents, are not shown. The solid line, a nonparametric regression smoother (Cleveland, 1979), summarizes the main trend of family size to increase with increasing cumulative family frequency, nearly linearly so in the bi-logarithmic plane for $N_f > 200$. In spite of this evident correlation, substantial variance in cumulative family frequency is also clearly visible for almost the full range of family sizes. If indeed family size is the crucial factor in Experiment 2, then the cumulative family frequency provides only an imperfect and noisy estimate of the variable that is really involved. Even in the absence of a correlation of the dependent variables in our experiments with the cumulative family frequency, a correlation with family size might still exist.

Interestingly, in the low cumulative family frequency condition more than half of the items do not have any morphological descendants. This suggests that the correlations observed for the low cumulative family frequency condition might in fact arise due to a nearly categorical distinction between zero family size ($N_f = V_f = 0$) and a small family size ($N_f > 0,\ V_f > 0$).

To further explore this possibility, we calculated the family size for the nouns appearing in Experiment 2. The high condition has a mean family size of 22.1 descendants (median 15.5), the low condition has a mean family size of 2.2 descendents (median 0). These numbers illustrate how serious the confound of cumulative family frequency with family size is. This confound receives further support from correlation analyses. Family size (in log units) is reliably correlated with reaction time in both the high condition (RT: $r = -.52, t(30) = -3.34, p < .002$) and in the low condition ($r = -0.50, t(30) = -3.13, p < .002$).

What these post-hoc analyses suggest is that the influence of the morphological family of a monomorphemic nouns on its recognition should be evaluated not in terms of tokens or some weighted token frequency count, but rather in terms of types. If correct, this hypothesis predicts that when we vary family size V_f, the type count, while keeping the token frequency counts, stem frequency and cumulative family frequency, constant, a significant difference should be observed. Conversely, no such difference should be observed when stem frequency and family size are kept constant, while varying the cumulative family frequency. These predictions are tested in Experiments 3 and 4.

Experiment 3

Experiment 3 investigates the role of family size in the processing of singular monomorphemic nouns matched for stem frequency and cumulative family frequency, as well as for various other variables. If indeed family size is an important factor in lexical processing, we should observe faster lexical decision times and higher subjective frequency ratings for nouns in the high condition (mean family size 20) than for nouns in the low condition (mean family size 4).

Method

Participants. Twenty-eight participants, mostly undergraduates at Nijmegen University, were paid to participate in the reaction time experiment, and forty different participants were paid to perform the subjective frequency experiment. All were native speakers of Dutch. None had participated in any of the preceding experiments.

Materials. We selected our word materials from the medium surface frequency range of approximately 80 to 550 occurrences per 42 million, using the CELEX lexical database. Thirty-six nouns were selected. Half of these nouns had a high family size (mean $V_f = 20.4$), the other half had a low family size (mean $V_f = 4.2$). We matched these two sets for the frequency of the singular form, the frequency of the plural form, cumulative family frequency, mean number of homonymic readings, as well as for the number of syllables, word length in letters, and geometric mean bigram frequency. In addition, 10 monomorphemic practice words were selected from the same frequency range as the target words. Forty-six phonotactically legal nonwords were constructed by changing one or two letters in monomorphemic words of the same frequency range.

Procedure. The procedure was completely identical to that of the preceding experiments.

Results and Discussion

For the lexical decision experiment, the data of two participants with an overall error score exceeding 10% were removed from the data set. As predicted, the nouns with a high number of descendents were responded to more quickly than the nouns with a low number of descendents, as shown in Table 3. For the reaction time data, the observed differences of 41 ms was highly significant, $F1(1, 25) = 38.2, p < .001$; $F2(1, 34) = 6.3, p < .02$.

The corresponding error percentages revealed a significant difference only by participants, $F1(1, 25) = 5.2, p < .05$; $F2(1, 34) = 2.2, p < .2$. For the subjective frequency rating, the difference of .93 units on a seven point scale was reliable both by participants and by items, $F1(1, 39) = 26.1, p < .001$; $F2(1, 34) = 8.9, p < .01$.

INSERT TABLE 3 APPROXIMATELY HERE

Given the potential importance of family size as a new factor in word recognition, we also calculated the mean number of orthographic neighbors for the two sets in this experiment, to make sure that the effect we have observed is not driven by neighborhood properties. (An orthographic neighbor is defined as a string of the same length with the same letters at the same positions except for one. Since the morphological descendants of monomorphemic nouns in Dutch are always obtained by affixation and hence of longer length, the set of neighbors and the morphological family of a given word are disjunct.) For the high condition, the mean number of neighbors was 4.3, for the low condition this number was 1.7. Note that this difference goes against our hypothesis and the observed direction of the difference in RT, since generally it is found that a higher number of neighbors slows down lexical processing (Luce, 1986; Grainger, O'Regan, Jacobs, & Segui, 1989; Grainger & Jacobs, 1996; but see Andrews, 1989). Furthermore, post-hoc correlations analyses did not reveal significant correlations (in one-tailed tests) of reaction time with number of neighbors ($r = .21, t(34) = 1.25, p > .1$), with mean neighbor frequency ($r = .14, t(34) = .82, p > .10$), nor with the number of higher-frequency neighbors ($r = .21, t(34) = 1.25, p > .10$).

We conclude that family size is a factor in visual word recognition that is independent of neighborhood density and neighborhood frequency. A difference in family size of 16 descendents is enough to give rise to a substantial difference in reaction time of 40 ms. This result suggests that the frequencies of the descendents of a monomorphemic noun have to be weighted on the basis of a type counts and not on the basis of their cumulated token frequencies. This hypothesis predicts that when we vary the cumulative family frequency while keeping all other factors, including family size, constant, no significant difference in reaction time or rating should appear. This prediction is tested in Experiment 4.

Experiment 4

Method

Participants. Thirty-three participants, mostly undergraduates at Nijmegen University, were paid to participate in the reaction time experiment, and forty different participants were paid to perform the subjective frequency experiment. All were native speakers of Dutch. None had participated in any of the preceding experiments.

Materials. We selected our word materials from the medium surface frequency range of approximately 28 to 545 occurrences per 42 million, using the CELEX lexical database. Thirty-four nouns were selected. Half of these nouns had a high cumulative family frequency (mean $N_f = 1007$), the other half had a low cumulative family frequency (mean $N_f = 38$). The two sets of nouns were matched for singular and plural frequency, for family size, as well as for bigram frequency, orthographic length in syllables and letters, and number of orthographic neighbors. Ten monomorphemic nouns were selected from the same frequency range as the target words to serve as practice items. Forty-four phonotactically legal nonwords were constructed by changing one or two letters in monomorphemic nouns from the same frequency range.

Procedure. The procedure was completely identical to that of the preceding experiments.

Results and Discussion

Ten participants performed with an error percentage greater than 10%. Their data were removed before further analyses. Two words were removed from the analysis due to an error score greater than 30%. Table 4 lists mean reaction time, error percentages, and mean subjective frequency ratings for the two experimental conditions. Analyses of variance revealed no significant effect of cumulative family frequency in the reaction time data, nor in the error data, nor in the ratings. In fact, all relevant F ratios were less than 1, except for the ratings in the by-participant analysis ($F(1,39) = 1.8, p > .15$).

INSERT TABLE 4 APPROXIMATELY HERE

Taken jointly, Experiments 3 and 4 show that the effect of morphological descendents on the processing of monomorphemic words is based on type frequency (family size), and not on token frequency (cumulative family frequency). Given that family size is an unexpected variable substantially influencing reaction times in lexical decision as well as subjective frequency ratings, we need to reconsider Experiment 1. Recall that Experiment 1 focused on the role of the frequency of the plural as a codeterminant of reaction time and rating of the singular. In Experiment 1, we did not control for possible effects of family size or cumulative family frequency. Given the results of Experiment 4, the fact that we did not match the materials of Experiment 1 for cumulative family frequency is irrelevant. However, the question remains to what extent family size appears as a confound in both Experiment 1 and in Experiment 2 of Baayen, Dijkstra, and Schreuder (1997). Before further investigating the nature of the family size effect, we first address this potential confound.

Reanalysis of the plural frequency effect

Experiment 1 revealed a significant effect of the frequency of the (unseen) plural form on the recognition latencies of monomorphemic singular nouns. In that experiment we did not control for variation in family size V_f. However, it turns out that for nouns in Dutch in general V_f is correlated with the frequency of the plural form ($rs = 0.49$). In our materials, we observed the same pattern ($rs = .614$). Given this substantial correlation, it is not surprising to find that (log) family size is strongly correlated with both reaction times ($r = -.64, t(68) = -6.96, p < .001$) and ratings ($r = .53, t(68) = 5.15, p < .001$) in Experiment 1. This raises the question of the relative importance of plural frequency on the one hand and family size on the other. To answer this question, we have run both parametric and nonparametric regression analyses.

A linear regression analysis with $\log(Fpl + 1)$ and $\log(V_f + 1)$ as independent variables and RT as dependent variable revealed a significant effect for family size ($t(67) = -4.85, p < .001$) but a marginally significant effect for plural frequency ($t(67) = -1.62, p < .06$). A nonparametric regression analysis using regression trees (Breiman, Friedman, Olshen, & Stone, 1984) similarly revealed family size to be the factor leading to the greatest reduction in the variance (more precisely, node heterogeneity). Further regression analyses of the subsets of words with a high versus a low plural frequency revealed that $\log(V_f + 1)$ is significantly correlated with reaction times in the low condition ($t(32) = -4.62, p < .0001$). Not surprisingly, $\log(Fpl + 1)$ did not reveal a significant effect in the low condition, the condition in which the plural frequency was kept as low as possible ($t(32) = -1.08, p > .10$). Crucially, $\log(Fpl+1)$ revealed a solid effect in the high condition ($t(32) = -3.56, p < .001$), and here the effect of family size is only marginally significant ($t(32) = -1.39, p < .09$). These analyses show that the frequency of the plural form is an independent factor in lexical processing, and that when the plural frequency is quite low the effect of family size is more strongly felt, but that when the plural frequency is high, the effect of the plural frequency is the primary factor.

We have also carried out post-hoc analyses of Experiment 2 in Baayen, Dijkstra, and Schreuder (1997), where we suspected that family size might similarly play an important role. Although log (V_f+1) by itself is indeed correlated with reaction time in this experiment ($r = -.48, t(67) = -4.44, p < .001$), both parametric and nonparametric regression analyses with the plural frequency and family size as independent variables and response latency as the dependent variable revealed that in this experiment the frequency of the plural (Fpl) is the only reliable predictor of reaction times ($t(66) = -3.42, p < .001$, for $\log(Fpl + 1)$; $t(66) = -0.44, p > .3$, for $\log(V_f + 1)$). We conclude that the frequency of the unseen plural form is indeed a reliable codeterminant of the processing speed of monomorphemic nouns.

Having addressed the correlation of family size with the frequency of the plural form, we are left with one other possible correlation, namely with the surface frequency of the monomorphemic singular form itself. For the nouns in Dutch in general, we indeed observe a large correlation ($rs = .61$) In the materials of Experiment 1, in which we contrasted nouns matched for singular frequency but with high versus low plural frequencies, the frequency of the singular form is also correlated with family size ($rs = .31$). Although (log) singular frequency (Fsg) is correlated with reaction time in a simple correlation ($r = -.29, t(68) = -2.47, p < .01$), we do not know to what extent this correlation is in fact due to the influence of family size. In a multiple regression analysis with $\log(V_f+1)$, $\log(Fsg+1)$, and $\log(Fpl+1)$ as independent variables and RT as dependent variable, the frequency of the singular does not appear as a significant factor ($t(66) = -.89, p > .15$), whereas the plural frequency and family size both emerge in a stepwise regression analysis as significant factors ($t(66) = -1.68, p < .05$ for the plural frequency; $t(66) = -4.30, p < .0001$ for family size).

Does this imply that surface frequency is irrelevant in the recognition of monomorphemic nouns in Dutch? To our mind, this is unlikely. We suspect that the influence of family size is primarily a more central, possibly semantic effect for which tasks such as lexical decision and subjective frequency rating are highly sensitive (for the sensitivity of lexical decision to the semantic properties of words, see, e.g., Balota, 1990; Coolen, Van Jaarsveld, & Schreuder, 1993; Jastrzembski, 1981; and Millis & Button, 1989). That is, we suspect that a given monomorphemic noun activates not only its own semantic representations, but also to a certain extent the semantic representations of its morphological relatives which, after all, are semantically quite similar. Conversely, the frequency of the singular form itself probably is primarily a form-based effect at the level of access representations. In Experiment 1, where we attempted to keep the singular frequency constant across conditions, the effect of singular frequency may well be masked by the substantial influence of family size. Hence we expect that when family size is kept constant as much as possible, we will still obtain a solid effect for singular frequency. Such an effect should not only be obtained with tasks such as lexical decision and subjective frequency rating, but also with progressive demasking.

In progressive demasking, a task developed by Grainger and Segui (1990), a word is presented on the screen by means of a continuous series of presentation cycles of equal duration. In the first cycle, the word is presented for 16 ms only, after which a series of hash marks is presented for 284 ms. With each successive cycle of 300 ms, the duration of the mask is decreased by a fixed amount of time (16 ms), so that the word is presented 16 ms longer. The impression for the participant is that a word is slowly emerging from a mist of hash marks. Initially, no word can be discerned at all, and it is only after a substantial decrease in masking that a word can be identified. Participants are asked to press a button as soon as they can identify a word. Response latencies typically vary between 1500 and 2200 milliseconds, depending on factors such as frequency and word length. These long latencies do not imply that participants have been seeing the word for that time. On the contrary, it takes participants such an amount of time to identify which word has been slowly emerging from the mist of hash marks. Progressive demasking, in other words, reduces the rate at which sensory information becomes available, thus slowing down the recognition process, and magnifying effects of visual identification processes. Thus, this is a task which primarily taps into the process of visual identification (Grainger & Segui, 1990; Grainger & Jacobs, 1996), and which is less sensitive to more central lexical processing (see also Van der Weide, 1997).

In Experiment 5 we therefore varied the frequency of the singular while matching for all other factors including family size, using lexical decision, subjective frequency rating, and progressive demasking as experimental tasks. If singular frequency is indeed a factor influencing perceptual identification, it should emerge in all three tasks. Moreover, if our intuition about the central nature of the effect of family size is correct, then the effect of family size should disappear when we replicate Experiment 3, which varied family size, but now using progressive demasking instead of visual lexical decision. This prediction is tested in Experiment 6.

Experiment 5

Method

Participants. Thirty participants, mostly undergraduates at Nijmegen university, were paid to participate in the lexical decision experiment. Forty different participants took part in the subjective frequency rating, and twenty-seven additional participants performed the progressive demasking task. None had participated in any of the preceding experiments.

Materials. Forty monomorphemic singular nouns were selected from the CELEX lexical database. Twenty nouns had a high singular frequency (mean: 543 per 42 million), the remaining twenty nouns had a low singular frequency (mean: 24 per 42 million). The two sets of nouns were matched for the frequency of the plural, family size and cumulative family frequency, the mean number of homonymic readings, and other potentially relevant factors. For the lexical decision task, ten monomorphemic nouns were selected from the same frequency range as the target words to serve as practice items. Fifty phonotactically legal nonwords were constructed by changing one or two letters in monomorphemic nouns from the same frequency range.

Procedure. For the rating and lexical decision tasks, the procedure was completely identical to that of the preceding experiments. Participants were tested individually in noise-proof experimentation booths in the progressive demasking task. The word stimuli were presented in alternation with a pattern mask consisting of a series of hash marks of equal length as the words themselves. On each successive cycle, the presentation of the word was increased by 16 ms, and the presentation of the mask was decreased by 16 ms. The total duration of each cycle remained constant at 300 ms. On the first cycle, the mask was presented for 284 ms, and the word for 16 ms. On the second cycle, the words were presented for 32 ms, etc. There was no interval between cycles. Cycles continued until the participant pressed the response key to indicate that she or he had recognized the word. The screen went blank after response initiation. Response latencies were measured from the beginning of the first cycle. Following response, participants were asked to write down the word they thought they had recognized.

Results and Discussion

This experiment led to somewhat higher error rates than in the previous experiments in the lexical decision task, a consequence of the very low frequencies of the words in the Low Frequency condition compared to the words in the other experiments. For the analyses of the reaction time experiment, we included participants with error scores less than 20% (The distribution of error scores in this experiment was shifted to the higher error region compared to our other experiments. However, all error scores below 20% fell within the bulk of the distribution. No outlier scores were included. A cutoff point at the 10% level was observed to give rise to the same pattern of significance, although with a reduced number of participants. Thus, the data of seven participants were removed from further analyses. Similarly, four nouns from the low singular frequency condition were removed due to error percentages exceeding 30%. Their removal did not affect the matching of the two data sets. Table 5 lists the mean reaction times and error percentages for the two conditions of this experiment.

As expected, a high singular frequency led to shorter response latencies than a low singular frequency. The 80 ms difference is significant both by participant and by item in reaction times and errors ($F1(1,22) = 113.8, p < .001$; $F2(1,34) = 19.7, p < .001$ for reaction times, $F1(1,22) = 12.4, p < .002$; $F2(1,34) = 5.9, p < .05$ for the error percentages). The subjective frequency ratings revealed the same pattern. Nouns with a high singular frequency were rated 1.5 units higher on a seven point scale than nouns with a low singular frequency. Again, this difference was significant both by participant and by item ($F1(1,39) = 91.4, p < .001$; $F2(1,38) = 17.9, p < .001$). Finally, the progressive demasking task showed the expected longer identification times for the nouns with a low singular frequency. The observed difference of 340 ms was significant both by participant and by item in reaction times ($F(1,26) = 107.7, p < .001$; $F2(1,38) = 19.1, p < .001$) and errors ($F1(1,26) = 9.45, p < .01$; $F2(1,38) = 6.36, p < .02$).

We conclude that the frequency of use of the singular form of monomorphemic nouns remains a solid factor in the early identification stages of word recognition. Although family size is correlated with singular frequency in general, singular frequency itself plays an independent role, even in the low frequency range (full range 0–35 per million, mean frequency in the high condition 13 per million, mean frequency in the low condition .6 per million) used in this experiment.

PLACE TABLE 5 APPROXIMATELY HERE

Experiment 6

The next experiment investigates whether family size is a factor in the early stages of form identification. Experiment 3 has shown that family size plays a substantial role in tasks such as lexical decision and subjective frequency rating, tasks which also tap into more central semantic processing. Experiment 5 showed that frequency effects pertaining to identification of form properties emerge in enlarged form in progressive demasking. If the effect of family size arises at later, more central stages of lexical access, i.e., after form identification, then no effect of family size should be observed for the word materials used in Experiment 3, which revealed an effect of family size in visual lexical decision, when used in a progressive demasking task. In other words, we expect a dissociation between lexical decision and subjective frequency ratings on the one hand and progressive demasking on the other, given that progressive demasking taps primarily into visual identification (Grainger & Segui, 1990; Grainger & Jacobs, 1996). We have observed a similar dissociation between progressive demasking and a decision task for Dutch compounds with linking phonemes. Identification times in progressive demasking are not affected by the spelling of the linking morphemes, while in contrast the decision task revealed a high sensitivity to these spelling changes, changes which affect the meanings of the compounds (Van der Weide, 1997). Our prediction that we will likewise observe a dissociation for the effect of family size is tested in Experiment 6.

Method

Twenty-nine participants, mostly undergraduates at Nijmegen university, were paid to participate in the identification experiment. All were native speakers of Dutch. None had participated in any of the preceding experiments. The word materials were completely identical to those used in Experiment 3 (see the Appendix). The procedure was identical to the progressive demasking procedure described for Experiment 5.

Results and Discussion

The mean identification times were 1840 ms for the high family size condition, and 1887 ms for the low family size condition (error percentages were 1.2 and 1.5 respectively). Analyses of variance across participants and items did not reveal any significant effects, neither for identification latencies nor for errors ($F < 1$ for all analyses). Since Experiment 5 has shown that progressive demasking is highly sensitive and can magnify a frequency effect obtained in lexical decision by a factor of three, the absence of any significant effect, let alone a magnified effect, argues against ascribing the nonresult to a lack of power or insensitivity of the progressive demasking task. Instead, we would like to argue that Experiment 6 provides evidence that morphological family size is a factor which operates at later stages of lexical processing.

This conclusion receives further support from a post-hoc analysis of the kind of morphologically complex words that enter into the counts of family size for the singular nouns used in Experiments 3 and 6. If the effect of family size is indeed a late, semantic effect, then we expect that the semantically transparent morphological descendents crucially drive the effect (see Marslen-Wilson, Tyler, Waksler, & Older, 1994, and Schreuder & Baayen, 1995, for the importance of semantic transparency in the processing of morphologically complex words). Inspection of the morphologically complex words involved in the counts of family size reveals that a small number of these descendents are semantically opaque. For instance, mafketel, "dumb-kettle," appears as a morphological descendent of ketel, "kettle," but has no clear semantic relation with its meaning, "stupid person." We have therefore compiled a second count of family size, now including only those morphologically complex words for which the meaning of the whole is clearly related to the meaning of the target noun. From a linguistic point of view, semantic transparency is a crucial condition for a word formation rule to be productive. From a psycholinguistic point of view, semantic transparency likewise plays a substantial role in lexical processing. Marslen-Wilson et al. (1994) have shown that morphological processing only takes place for semantically transparent formations. Thus, we expect that when we enhance our counts by excluding semantically opaque words, the correlation with the reaction times should improve. Although somewhat impressionistic, our enhanced count of the family size, V_f', indeed revealed a higher correlation with the reaction times of Experiment 3. For the raw count, the correlation of $\log(V_f + 1)$ with RT was $r = -.38$ ($t(34) = -2.40, p < .02$), for the enhanced count, the correlation of $\log(V_f' + 1)$ with RT was $r = -0.42$ ($t(34) = -2.79, p < .01$). This increase in the correlation suggests informally that semantic transparency might play a role. Clearly, further experimental investigation is required to shed further light on this issue. Nevertheless, in combination with the absence of an effect of family size in progressive demasking, the indication that semantic transparency might be at issue suggests that in all likelihood family size exerts its effect at later, more central stages of lexical processing.

General Discussion

In this study we have focussed on three properties that codetermine the speed with which monomorphemic nouns in Dutch are processed: the frequency of the (uninflected) singular form itself (Experiment 5), the frequency of the corresponding plural form (Experiment 1), and the number of morphologically complex words that contain that particular noun as one of its constituents (Experiments 2 and 3). We have found that this number of morphologically related words, what we have called the morphological family size, is a strong independent determinant of both response times in lexical decision as well as of subjective frequency ratings. In contrast, the cumulative family frequency, the summed token frequencies of the words in the morphological family, plays no role at all (Experiment 4).

Progressive demasking is a task that primarily captures aspects of early identification stages of lexical processing (Grainger & Segui, 1990; Grainger & Jacobs, 1996). In order to ascertain whether the effect of family size occurs relatively early or late, we used the same materials that yielded a solid effect in lexical decision and subjective frequency rating (Experiment 3) in a progressive demasking task (Experiment 6). In Experiment 6, we did not observe any effect of family size. In contrast, the effect of the frequency of the singular form appears as an 80 ms effect in lexical decision and as a 340 ms effect in progressive demasking (Experiment 5). We therefore hypothesize that the effect of singular frequency is already present in the early stages of perceptual identification, while the effect of family size probably arises following perceptual identification.

The unexpected emergence of morphological family size as a factor in the lexical processing of monomorphemic nouns raises several issues. Before discussing these in some detail, we first comment on a question of methodology. Obviously, family size and plural frequency are factors that should be taken into account when performing experiments with monomorphemic nouns. At the very least, variation in family size may introduce substantial noise in reaction time data. More interesting is the following issue. Gernsbacher (1984) has pointed out that frequency counts, especially in the lower frequency ranges, are subject to sampling error, and that for words in these lower frequency ranges subjective frequency ratings may be more reliable than frequency counts. Due to sampling error, especially in small corpora, frequencies of use in the lower frequency ranges may be underestimated. Interestingly, Experiment 5 shows that even in the lower frequency ranges (1–14 per million) one may obtain very reliable and solid frequency effects when less noisy and more reliable frequency counts are used that are based on large corpora. In this experiment, log singular frequency (using the CELEX frequency counts for a 42 million word corpus) and reaction time revealed a robust effect, not only in the factorial design, but also in a post-hoc correlation analysis ($r = -0.53, t(34) = -3.65, p < .001$). Furthermore, the subjective frequency ratings likewise reveal a high correlation with singular frequency (for log (Fsg+1) and Rating, $r = 0.43, t(34) = 2.81, p < 0.01$). This shows that low frequencies of use, conditional on their being adequately sampled, may reveal contrasts that are consistently present in both lexical decision reaction times and in subjective frequency ratings. (Even with a large corpus such as that analyzed in the CELEX lexical database, obvious sampling errors remain. For instance, friet, "French fries," an extremely popular kind of junk food in the Netherlands, has a frequency of only 2 per million in our counts. In this particular example, the sampling error arises due to the exclusion of spoken Dutch in the corpus. We excluded from our experimental lists items with such glaring mismatches between frequency in CELEX and obvious everyday frequency in our daily experience.)

In fact, throughout our experiments the subjective frequency ratings reveal correlations with lexical decision times that are stronger than any other factor studied here. For instance, in Experiment 5, which focused on the role of singular frequency, ratings and RT in lexical decision have a correlation of $r = -.75$ ($t(34) = -6.55, p < .001$), whereas RT and log (Fsg+1) have a somewhat lower correlation ($r = -0.53, t(34) = -3.65, p < .001$). The consistency in the results obtained by means of lexical decision and those obtained by means of subjective frequency ratings suggests that with respect to issues of type and token frequencies the two tasks tap into similar aspects of lexical organization, and that therefore in this domain of inquiry subjective frequency rating is an excellent pretest and perhaps a (cheaper and faster) alternative for lexical decision. Note, however, that subjective frequency ratings do not always reliably predict progressive demasking identification times. In Experiment 5, the ratings are still reliably correlated with identification times ($r = -0.48, t(34) = -3.18, p < 0.01$, one-tailed test), albeit to a lesser extent than with the response latencies in lexical decision ($r = -0.75, t(34) = -6.55, p < .001$). This suggests that the ratings pick up aspects of form familiarity. But the dissociation between identification times and ratings observed for Experiments 3 and 6 shows that the ratings also pick up aspects of later central processing occurring after perceptual identification. Thus subjective frequency ratings appear to be sensitive to two aspects of lexical familiarity: form and meaning.

But what aspects of meaning are picked up by means of family size in lexical decision latencies and subjective frequency ratings? How should we explain this effect of family size? The model of morphological processing developed in Schreuder and Baayen (1995) and Baayen, Dijkstra, and Schreuder (1997) provides a theoretical framework within which the effect of family size can be understood. According to this model, form-based access representations map onto lemma nodes that in turn map onto syntactic and semantic representations. The lemma nodes crucially link form information at the access level with higher-order semantic and syntactic information. Semantically similar words share meaning representations, that is, at least some of the links from their lemma nodes to the semantic layer lead to the same representations. The model has two flows of activation, a forward flow from form to meaning, and a backward flow from meaning to form. For instance, when reading dog, the corresponding access representation activates the lemma node for dog, which in turn activates the semantic representations for canines as well as the appropriate syntactic representations for number and word category. Crucially, the model assumes that the access representation of dog only activates its own lemma node. However, due to the subsets of shared semantic representations, related lemma nodes for words such as doghouse, dogfight, doggy, and dogcart also become activated during the backward flow of activation from the semantic representations to the lemma nodes, in the same way as table may become partially activated upon reading chair. Thus, presentation of a word with a large morphological family leads to the activation of a large number of lemma nodes. Words with few morphological descendents, to the contrary, activate only a few lemma nodes. It is this difference in the number of activated lemmas, or, more precisely, the logarithm of this number, that we suspect to give rise to higher ratings and shorter lexical decision latencies (see Grainger & Jacobs, 1996, for a computational model that takes the monitoring of global lexical activation in the lexical decision task into account). Further empirical investigation is clearly required to further substantiate this tentative explanation.

The model as outlined in Schreuder and Baayen (1995) claims that in the backward pass of activation, the resting activation levels of the access representations of the constituents of a complex word that has been encountered are slightly increased. This increase is argued to take place both for inflected words such as dogs as well as for compounds and derivations. Hence, the model predicts both an effect of plural frequency and an effect of cumulative family frequency. Experiment 4, however, shows that there is no effect of cumulative family frequency, which falsifies the idea that the backward flow of activation to the access representations takes place for derivations and compounds. Experiment 1, by contrast, shows that the frequency of the plural form is a crucial determinant of response latencies in visual lexical decision and subjective frequency ratings. In Schreuder and Baayen (1995), we hypothesized that the effect of plural frequency on the processing of the singular takes place at the level of access representations due to the backward flow of activation. However, alternative explanations are equally well possible at more central levels of lexical processing. We are currently carrying out a series of experiments in which we investigate these alternative possibilities in detail (see also Baayen, Lieber, & Schreuder, 1997). For the present discussion, the main observation that we have to offer is the dissociation between the plural frequency and cumulative family frequency effects.

Note that this dissociation — a cumulative token frequency effect for nouns with respect to their plurals, but no such effect for derived words and compounds — coincides with the distinction between inflectional and derivational morphology. The distinction between inflection on the one hand and derivation and compounding on the other (Lyons, 1968, p. 195) is intimately linked with a difference in the kind of semantic operation involved. Inflectional semantic operations are generally quite regular and predictable, derivational operations and compounding almost always add unpredictable shades of meaning that cannot be deduced from the meanings of the constituents in isolation (Lyons, 1977, p. 524). Such semantic idiosyncrasies must lead to extensive storage of the meanings of derived words and compounds in the mental lexicon. Thus many derived words and compounds probably have autonomous lexical entries, entries which themselves participate in the inflectional paradigms of the language. The substantial effect of family size provides evidence that indeed the meanings of large numbers of complex words are stored. Apparently, the constituents of autonomous complex words do not benefit at the access level from repeated exposure to these complex words. It is only for semantically completely predictable complex words, i.e., inflected words, that we can observe a cumulative frequency effect on the base word. It is surprising that the different properties of inflectional and derivational morphology can be detected by studying the processing of monomorphemic nouns.

Given the presence of a family size effect and the absence of a cumulative family frequency effect for monomorphemic nouns, the question arises whether a similar state of affairs holds for inflected and derived words as well as for compounds. Here we are faced with a doubling of complexities, as we now have to take into account the frequency counts of both the complex word itself and those of its constituents. With respect to a complex word such as fisher, we may expect effects of its own surface frequency, its stem frequency (the summed frequencies of its own inflectional variants fisher and fishers), and its family size (the number of words derived from fisher, such as fisherman, fishery, pearl-fishery, etc.). Given our present results, we do not expect any effect of the cumulative family frequency of fisher.

In models in which on-line morphological parsing can take place, the constituents of a word, such as fish and -er in fisher, may also play a role in the recognition process. Again, our findings suggest that the most probable potential factors are the stem frequency of fish, the family size of fish, but not the cumulative family frequency of fish.

In the literature, the possible role of stem frequency and family size have not been investigated for complex words. In contrast, several researchers have addressed the possible role of the frequency properties of the constituents of complex words. In the domain of inflectional morphology, Taft (1979), Burani et al. (1984) and Burani and Caramazza (1987) contrasted inflected words with a high and low stem frequency of the base word. Typically, when matched for their own frequency of use, inflected words with a high stem frequency of the base were responded to more quickly in visual lexical decision than words with a low stem frequency of the base. Although these authors disagree about how these results should be modeled, they take these results to indicate that stem and affix representations play a functional role during the recognition of inflected words. In addition, Taft (1979, Experiment 2) also pointed out that uninflected monomorphemic words with a high stem frequency are processed more quickly in lexical decision than uninflected monomorphemic words with similar surface frequency but with a low stem frequency. Even though their results might be influenced by the correlation between the stem frequency of the base and the family size of the base, their general conclusions are presumably still valid given the results of our Experiment 1.

In the domain of derivational morphology and compounding, Taft and Forster (1976), Taft (1979), Bradley (1979), and Cole et al. (1989) have similarly investigated the possible role of the constituents. Taft and Forster (1976, Experiment 5) contrasted compounds of equal surface frequency with respect to the surface frequency of the first constituent. They found that compounds with a high-frequency first constituent elicited shorter response latencies in lexical decision than compounds with a low-frequency first constituent. They concluded that such compounds are recognized via the first constituent. Taft and Forster are among the few who report not only their materials but also the mean reaction times for their experimental words, as should be rule rather than exception for these kinds of experiments. Thus we can re-analyze their results using the frequency counts in the CELEX lexical database. The mean frequencies of the first constituents for the two conditions are 24849 and 198 per 18 million. However, looking at the surface frequencies of the compounds, we found that the mean frequency in the high condition, 181, was nearly twice the mean frequency in the low condition (99). In addition, the first constituent in the high condition has a mean family size of 33, while the first constituent in the low condition has a mean family size of only 4. Post-hoc correlational analyses reveal no significant correlation of the log frequency of the first constituent with reaction time ($r = -0.09, t(38) = -0.53, p > 0.20$), but a significant correlation of log family size with reaction time ($r = -0.28, t(38) = -1.77, p < .05$), as well as a significant correlation of the log surface frequency of the compound with reaction time ($r = -0.47, t(38) = -3.25, p < .01$). To our mind, these analyses suggest that these compounds are processed on the basis of their full forms. Since our Experiments 3 and 6 suggest that the effect of family size is one that takes place at later stages of lexical processing, after perceptual identification, the effect of family size observed for Taft and Forster's compounds is unlikely to indicate that these compounds are identified by first accessing the initial constituent. The high correlation of log surface frequency with reaction time points in the same direction.

Independent evidence suggesting that family size may play a role in the processing of compounds can be found in Van Jaarsveld, Coolen, and Schreuder (1994). These authors found that novel compounds for which the summed family size of the first and second constituents is high were harder to reject as existing words than novel compounds with a low summed family size. Apparently, when a novel compound has constituents with large morphological families, the activation that spreads to these family members renders these novel compounds very word-like, and hence difficult to reject as an existing word.

Taft (1979), Bradley (1979), and Cole et al. (1989) investigated the role of the constituents of derived words by varying the sum of the stem frequency and the cumulative family frequency. Their experiments revealed an effect of cumulative token frequency: Words with a high cumulative token frequency are responded to more quickly than words with low cumulative token frequency counts. These results have led Schreuder and Baayen (1995) to model both stem frequency and cumulative family frequency by means of the same mechanism of activation feedback. We now suspect that when stem frequency and cumulative family frequency are carefully distinguished, and when family size is also taken into account, no effect of cumulative family frequency will be observed.

Taft (1979) reports a cumulative frequency effect for semantically opaque complex words with bound roots of Latin origin such as -vade in pervade and -plex in perplex. Given the results of Marslen-Wilson et al. (1994), which suggest that the constituents of semantically opaque words do not play a role in lexical processing, it is unclear to us whether this is a true morphological effect. To our mind, it is highly unlikely that putative stems such as -plex and -vade play any role in the mental lexicon (see Schreuder & Baayen, 1994), and we therefore suspect that substring familiarity at a pre-lexical level is at issue.

Complementary to these studies, Burani, Dovetto, Thornton, and Laudanna (1996) focussed on the frequency properties of affixes instead of stems. They argue that both family size and cumulative family frequency, factors that they distinguish but do not contrast factorially in their experiments, are exponents of the salience of affixes and as such codetermine response latencies in visual lexical decision and word naming. Again, on closer examination it may well turn out that the family size and not the cumulative family frequency is the driving factor. Obviously, further research on the processing of complex words and the role of family size therein is necessary. Our present results are suggestive but not as yet conclusive.

An inspection of the kind of words that appear in the counts for family size of our Dutch materials shows that compounding, the most productive morphological process in Dutch, is responsible for the bulk of the family members. For a language such as German, in which compounding is even more productive than in Dutch, we therefore also expect substantial effects of family size. With respect to English, the orthographic convention of writing many compound words with intervening spaces would at first sight suggest that perhaps family size is a less important factor in this language, even though compounding is quite productive in English. However, counts of family size and cumulative family frequency for English complex words (including compounds written without intervening spaces) nevertheless reveal the same kind of correlational structure shown in Figure 1 for Dutch. Moreover, a replication study using subjective frequency ratings for comparable English materials using native speakers of English revealed exactly the same pattern of results as the ratings reported here for Dutch (see Baayen, Lieber, & Schreuder, 1997). This suggests that family size may also be an important factor in the lexical processing of English monomorphemic words. For languages such as French, in which compounding is much less productive, it is unclear whether family size, defined in terms of the number of related complex words, will turn out to be equally important as in Dutch or English. But it is possible that our definition of family size is too restrictive, and that counts of semantically transparent fixed phrases such as chemin de fer, "railroad," and flûte traversière, "flute," which denote well-known concepts, should also be included.

In sum, for singular nouns the token frequencies of the singular form as well as the token frequencies of the plural form codetermine lexical processing in visual word recognition. At present, it is unclear whether the plural token frequency effect occurs early at the level of form identification, or whether it is a relatively late, more central phenomenon. Family frequency, by contrast, does not appear to affect lexical processing at all, while family size emerges as a factor influencing lexical processing after form identification. Family size appears to be an indicator of the extent to which a noun is incorporated in the network of semantic relations linking concepts in the mental lexicon. By definition, family frequency is a type frequency count. As such, its effect on lexical processing can only be explained under the assumption that many complex words have their own semantic representations in the lexicon. Without such representations, in a theory in which the meanings of complex words have to be computed on-line, it is difficult to see how a type-frequency effect could arise.

Evidently, the complexity of simplex words is such that a naive view of these words as lexical islands is untenable. The way in which simplex nouns are processed in various tasks can only be understood within a comprehensive approach to the mental lexicon, in which the structural, the semantic, the morphological, and the distributional properties of words are taken into account.

References

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

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

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual route model (submitted).

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database (CD-ROM), Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

Baayen, R. H., Lieber, R., & Schreuder, R. (1997). The morphological complexity of simplex nouns (submitted).

Balota, D. A. (1990). The role of meaning in word recognition. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), Comprehension processes in reading (pp. 9–32). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Bradley, D. C. (1979). Lexical representation of derivational relation. In M. Aronoff & M. L. Kean (Eds.), Juncture (pp. 37–55). Saratoga: Anma Libri.

Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). Classification and regression trees. Belmont, California: Wadsworth International Group.

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

Burani, C., Dovetto, F. M., Thornton, A. M., & Laudanna, A. (1996). Accessing and naming suffixed pseudo-words. In G. E. Booij & J. van Marle (Eds.), Yearbook of morphology 1996 (pp. 55–72). Dordrecht: Kluwer Academic Publishers.

Burani, C., Salmaso, D., & Caramazza, A. (1984). Morphological structure and lexical access. Visible Language,18, 342–352.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association, 74, 829–836.

Cole, P., Beauvillain, C., & Segui, J. (1989). On the representation and processing of prefixed and suffixed derived words: A differential frequency effect. Journal of Memory and Language, 28, 1–13.

Coolen, R., Van Jaarsveld, H. J., & Schreuder, R. (1993). Processing novel compounds: Evidence for interactive meaning activation of ambiguous nouns. Memory and Cognition,21, 235–246.

Gardner, M. K., Rothkopf, E. Z., Lapan, R., & Lafferty, T. (1987). The word frequency effect in lexical decision: Finding a frequency-based component. Memory & Cognition,15, 24–28.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. Journal of Experimental Psychology: General,113, 256–281.

Gordon, B. (1983). Lexical access and lexical decision: Mechanisms of frequency sensitivity. Journal of Verbal Learning and Verbal Behavior,22, 24–44.

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

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing word units in visual word recognition: The neighborhood frequency effect. Perception & Psychophysics,45, 189–195.

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

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. Perception & Psychophysics, 28, 267–283.

Hasher, L. & Zacks, R. T. (1984). Automatic processing of fundamental information. the case of frequency of occurrence. American Psychologist, 39, 1372–1388.

Jastrzembski, J. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. Cognitive Psychology, 13, 278–305.

Jescheniak, J. D. & 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. & Burani, C. (1985). Address mechanisms to decomposed lexical entries. Linguistics, 23, 775–792.

Luce, P. A. (1986). Neighborhoods of words in the mental lexicon. Research on speech perception technical report 6. Bloomington: Indiana University.

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

Lyons, J. (1977). Semantics 2. Cambridge: Cambridge University Press.

Marslen-Wilson, W., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. Psychological Review, 101, 3–33.

Millis, M. L. & Button, S. B. (1989) The effect of polysemy on lexical decision time: Now you see it, now you don't. Memory & Cognition, 17, 141–147.

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

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977) Frequency and repetition effects in lexical memory. Journal of Experimental Psychology: Human Perception and Performance, 3, 1–17.

Schreuder, R. & Baayen, R. H. (1994). Prefix-stripping re-revisited. Journal of Memory and Language, 33, 357–375.

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.

Shapiro, B. J. (1969). The subjective estimation of word frequency. Journal of Verbal Learning and Verbal Behavior, 8, 248–251.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. Memory & Cognition, 7, 263–272.

Taft, M. & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. Journal of Verbal Learning and Verbal Behavior, 15, 607–620.

Van der Weide, F. (1997). Visuele identificatie en lexicale verwerking van samenstellingen met bindfonemen (Visual identification and lexical processing of compounds with linking phonemes). Unpublished Master's Thesis, University of Nijmegen, Nijmegen.

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

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

Author Note

R. Schreuder, Interfaculty Research Unit for Language and Speech, University of Nijmegen, The Netherlands. R.H.Baayen, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

We are indebted to Leonore Biegstraten for her aid in running the experiments, and to James McQueen and Anne Cutler for invaluable discussion. We would also like to express our gratitude to three reviewers and to Gregory Murphy for their helpful comments. Correspondence concerning this article should be addressed to R. Schreuder, Interfaculty Research Unit for Language and Speech, University of Nijmegen, Wundtlaan 1, 6525 XD, Nijmegen, The Netherlands. Electronic mail: rob.schreuder@mpi.nl.

Appendix

Materials used in the experiments. After each word and its English translation, we list mean RT in visual lexical decision, mean subjective frequency rating, and, if applicable, mean identification time in progressive demasking. Words highlighted with a † symbol were not considered in the data analysis due to high error rates.

Words used in Experiment 1.

Nouns with a low plural frequency: bast (bark): 663, 1.70; blos (bloom): 614, 2.63; bries (breeze): 590, 2.52; buit (booty): 594, 2.26; dij (thigh): 596, 3.11; doorn (thorn): 588, 2.33; dwang (compulsion): 519, 2.93; gif (poison): 551, 3.07; gips (plaster): 555, 2.48; glimp (glimpse): 629, 1.93; greppel (channel): 627, 1.78; griep (flu): 537, 4.93; ham (ham): 557, 4.56; haver (oats): 501, 2.22; honing (honey): 535, 3.23; juf (teacher): 577, 4.52; kachel (stove): 521, 4.41; kraag (collar): 527, 3.96; mos (moss): 590, 2.30; nonsens (nonsense): 628, 3.63; oven (oven): 557, 4.67; panter (panther): 549, 2.26; pasta (paste): 524, 4.89; peer (pear): 500, 3.96; podium (stage): 549, 2.81; puin (rubble): 604, 2.63; reuma (rheumatism): 642, 2.56; saldo (balance): 600, 3.48; schuit (barge): 598, 2.00; stier (bull): 502, 3.59; sultan (sultan): 608, 1.26; unie (union): 608, 2.67; weelde (luxury): 569, 2.07; wrok (resentment): 627, 2.04; zweem (trace): 698, 1.26.

Nouns with a high plural frequency: akker (field): 557, 2.96; boon (bean): 597, 3.56; cijfer (digit): 552, 5.48; deken (blanket): 519, 5.11; detail (detail): 549, 3.81; druif (grape): 573, 3.93; gans (goose): 609, 2.48; graad (degree): 558, 3.37; hiel (heel): 582, 2.78; insekt (insect): 537, 4.04; juweel (jewel): 529, 2.85; kruid (herb): 520, 3.52; laars (boot): 495, 3.78; les (lesson): 557, 5.48; letter (letter): 493, 5.48; lip (lip): 476, 4.70; luik (hatch): 530, 2.15; nagel (nail): 501, 4.81; oksel (armpit): 578, 2.89; parel (pearl): 545, 2.04; rat (rat): 533, 2.70; schaap (sheep): 507, 4.04; schoen (shoe): 527, 5.74; spier (muscle): 486, 4.89; spul (gear): 545, 5.11; stengel (stalk): 626, 2.37; struik (bush): 506, 4.11; tand (tooth): 499, 5.00; teen (toe): 549, 4.52; teugel (rein): 625, 2.04; ui (onion): 533, 5.15; vlam (flame): 499, 4.52; vrucht (fruit): 544, 4.70; wiel (wheel): 510, 4.59; wortel (root): 509, 4.33.

Mean frequencies for the two conditions: singular frequency: 309, 308; plural frequency:

14, 910; cumulative family frequency: 323, 539; morphological family size: 7.6, 22.9.

Words used in Experiment 2.

Nouns with a low cumulative family frequency: allure (airs): 646, 3.45; azijn (vinegar): 560, 6.42; banier (banner): 627, 2.06; bekken (basin): 626, 4.81; brancard (stretcher): 622, 5.23; brink (farmyard): 744, 2.23; burcht (castle): 660, 5.19; deeg (dough): 573, 6.48; geste (gesture): 785, 2.45; grind (gravel): 654, 6.03; hulst (holly): 635, 4.61; jade (jade): 605, 2.58; jeep (jeep): 661, 5.55; kanjer (whopper): 605, 6.03; karma (karma): 747, 2.03; klepel (clapper): 843, 4.74; kou (cold): 531, 6.74; maizena (corn flour): 802, 4.42; meute (pack): 677, 4.42; motel (motel): 654, 5.35; notie (notion): 540, 4.06; panda (panda): 649, 5.77; rum (rum): 629, 6.06; scala (scale): 615, 3.39; sherry (sherry): 590, 5.65; stress (stress): 595, 5.94; teneur (tenor): 716, 2.71; terp (terp): 685, 3.65; tijm (thyme): 626, 4.03; tred (step): 728, 3.71; venijn (poison): 686, 4.32; zone (zone): 568, 4.77; zwaard (sword): 582, 6.58.

Nouns with a high cumulative family frequency: alarm (alarm): 536, 6.58; ambacht (trade): 588, 5.71; ambt (office): 619, 5.06; arrest (arrest): 620, 5.19; dwang (compulsion): 516, 6.42; fatsoen (decorum): 594, 6.42; gigant (giant): 577, 4.94; ijzer (iron): 541, 6.68; kanton (canton): 726, 2.61; kerst (Christmas): 571, 6.90; kolos (colossus): 629, 4.03; kramp (cramp): 589, 6.32; mars (march): 586, 5.97; nijd (envy): 632, 5.32; pracht (magnificence): 666, 5.74; pret (fun): 510, 6.68; ratio (reason): 681, 3.32; sap (juice): 589, 6.74; schicht (flash): 726, 3.26; schild (shield): 564, 5.87; schrift (writing): 599, 6.81; soja (bean sauce): 648, 3.81; spil (pivot): 586, 4.45; spion (spy): 562, 6.13; spraak (speech): 508, 6.26; vaat (washing-up): 574, 6.16; vloed (tide): 552, 6.32; weelde (luxury): 595, 4.97; wol (wool): 522, 6.61; zege (victory): 606, 5.19; zicht (sight): 603, 6.48;

Mean frequencies for the two conditions: singular frequency: 253, 246; plural frequency: 23, 27; cumulative family frequency: 20, 2680; morphological family size: 2.2, 22.1.

Words used in Experiments 3 and 6.

Nouns with a low morphological family size: barbaar (barbarian): 591, 3.22, 1886; beton (concrete): 553, 4.78, 1631; faam (reputation): 628, 2.29, 1660; flank (flank): 643, 2.39, 2002;

fluweel (velvet): 524, 3.63, 1697; heiden (heathen): 692, 2.68, 2171; karton (cardboard): 552, 4.95, 1735; list (trick): 562, 3.22, 2308; ratio (reason): 689, 3.54, 2054; rebel (rebel): 639, 3.63, 2101; regio (region): 532, 4.93, 1733; rund (cow): 533, 4.51, 1894; scheut (twinge): 644, 3.10, 1916; smart (sorrow): 588, 2.83, 1986; tiran (tyrant): 596, 3.51, 1736; vulkaan (volcano): 513, 3.98, 1711; waas (haze): 639, 2.78, 1805; weelde (luxury): 581, 3.10, 2023.

Nouns with a high morphological family size: alarm (alarm): 518, 5.41, 1850; gif (poison): 576, 4.80, 1606; graan (grain): 523, 4.37, 1863; haag (hedge): 646, 2.59, 1491; kanon (gun): 586, 3.39, 1879; ketel (kettle): 548, 5.02, 1860; knol (tuber): 582, 3.37, 1974; koren (corn): 548, 3.76, 1896; kostuum (suit): 527, 4.02, 1553; meel (flour): 557, 4.83, 1962; rente (interest): 519, 5.85, 1872; rijst (rice): 516, 6.00, 1869; schijf (disc): 533, 5.00, 1833; taart (cake): 480, 5.85, 1727; tucht (discipline): 623, 2.49, 2151; vaat (washing-up): 610, 4.15, 2055; vracht (freight): 540, 5.07, 1837; worm (worm): 524, 4.15, 1714.

Mean frequencies for the two conditions: singular frequency: 193, 230; plural frequency: 84, 60; cumulative family frequency: 418, 404; morphological family size: 4.2, 20.0.

Words used in Experiment 4

Nouns with a low morphological family frequency: berk (birch): 686, 3.59; fuik (bow net): 645, 2.56; gazon (lawn): 594, 4.41; gong (gong): 675, 2.61; harp (harp): 568, 3.34; huls (case): 658, 2.63; kapel (chapel): 670, 3.80; koker (case): 629, 3.56; kreeft (lobster): 558, 3.83; lus (loop): 599, 4.05; merrie (mare): 719, 3.05; paneel (panel): 566, 2.90; schuit (barge): 642, 3.05; sofa (sofa): 608, 3.15; trog (trough)†: 726, 1.76; veen (peat): 666, 2.71; volume (volume): 542, 5.41.

Nouns with a high morphological family frequency: barbaar (barbarian): 593, 3.49; faam (reputation): 653, 2.51; fatsoen (decorum): 591, 4.78; fluweel (velvet): 531, 3.61; gigant (giant): 575, 3.78; graat (fish bone): 672, 3.61; heiden (heathen): 724, 2.66; kanton (canton): 682, 2.07; kommer (destitution)†: 862, 2.22; nijd (envy): 706, 3.22; pracht (magnificence): 584, 3.73; pret (fun): 511, 5.54; ratio (reason): 692, 3.54; regio (region): 566, 5.12; spil (pivot): 575, 3.12; unie (union): 586, 4.61; waas (haze): 638, 3.12.

Mean frequencies for the two conditions: singular frequency: 181, 182; plural frequency: 43, 41; cumulative family frequency: 38, 1007; morphological family size: 5.9, 5.8.

Words used in Experiment 5

Nouns with a low singular frequency: bivak (bivouac): 641, 3.22, 1736; bretel (braces): 757, 3.66, 2063; buks (rifle)†: 750, 3.73, 1869; fort (fort): 617, 4.20, 1669; frase (phrase): 707, 2.88, 2167; fregat (frigate): 731, 2.93, 1956; gondel (gondola): 659, 3.22, 2028; kaf (chaff): 745, 2.54, 2189; kieuw (gill): 664, 3.59, 1960; leeuwerik (lark): 590, 4.05, 2139; mammoet (mammoth): 646, 3.24, 1980; naaf (hub)†: 721, 1.90, 2456; pukkel (pimple): 569, 5.80, 2136; schalk (rogue)†: 663, 1.29, 2806; sering (lilac)†: 803, 3.00, 2010; spijl (bar): 604, 3.49, 2384; stola (stole): 697, 2.22, 2002; stronk (stump): 725, 4.00, 2136; ventiel (valve): 601, 5.41, 1885; vrek (miser): 597, 3.83, 2012.

Nouns with a high singular frequency: baai (bay): 576, 4.61, 2137; buit (booty): 598, 5.29, 1990; dozijn (dozen): 550, 5.51, 1747; gade (spouse): 724, 1.93, 1928; gazon (lawn): 547, 4.93, 1246; kade (quay): 554, 4.93, 1465; karwei (job): 565, 5.90, 1714; korporaal (corporal): 628, 3.71, 2018; ober (waiter): 565, 6.41, 1718; podium (apron): 540, 5.88, 1441; pond (pound): 541, 5.98, 1527; prooi (prey): 524, 5.61, 1713; ravijn (ravine): 534, 5.02, 1581; romp (trunk): 553, 4.49, 1539; sekte (sect): 583, 5.07, 1980; sofa (sofa): 570, 3.90, 1481; stank (stench): 636, 6.49, 1778; sultan (sultan): 585, 3.34, 1792; vijver (pond): 516, 6.22, 1712; wier (algae): 652, 3.24, 2157.

Mean frequencies for the two conditions: singular frequency: 24, 534; plural frequency: 30, 37; cumulative family frequency: 23, 26; morphological family size: 1.2, 1.3.

Table 1

Mean response latencies, error percentages (visual lexical decision), and subjective frequency ratings for monomorphemic singulars with high and low summed frequencies of their inflectional variants (Experiment 1).

	RT	Error	Rating
High summed frequency	539	2.3	3.95
Low summed frequency	580	3.2	2.91

Table 2.

Mean response latencies, error percentages (visual lexical decision), and subjective frequency ratings for monomorphemic singulars with high and low summed frequencies of morphologically derived words (Experiment 2).

	RT	Error	Rating
High summed frequency	599	2.6	5.54
Low summed frequency	644	11.8	4.60

Table 3.

Mean response latencies, error percentages (visual lexical decision), and subjective frequency ratings for monomorphemic singulars with high and low numbers of morphological descendents (V_f) (Experiment 3).

	RT	Error	Rating
High V_f	553	0.9	4.46
Low V_f	594	2.4	3.53

Table 4.

Mean response latencies, error percentages (visual lexical decision), and subjective frequency ratings for monomorphemic singulars with high and low cumulative descendent token frequency (N_f) (Experiment 4).

	RT	Error	Rating
High N_f	598	3.0	3.59
Low N_f	612	2.1	3.33

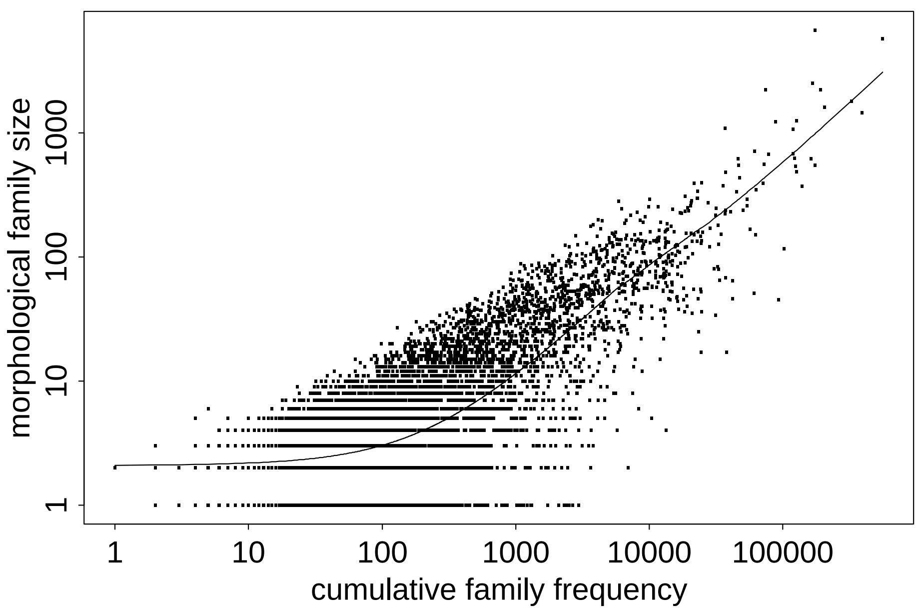
Table 5.

Mean response latencies, error percentages (visual lexical decision), subjective frequency ratings, and identification times (progressive demasking) for monomorphemic singulars with high and low singular frequency (Experiment 5)

	RT (Error)	Rating	Identification (Error)
High Fsg	576 (2.2)	4.93	1733 (.002)
Low Fsg	656 (7.5)	3.43	2073 (.024)

Figure Caption

Figure 1. Logarithmic scatterplot of the number of morphologically related tokens (N_f) and types (V_f) for Dutch morphemes. The solid line represents a nonparametric regression smoother (Cleveland, 1979). The morphemes plotted in the upper right of the graph are typically affixes. The morpheme with the highest number of descendents that is not an affix is the noun/verb stem *werk*, "work" ($V_f = 550$).


morphological family size
1
10
100
1000
1
10
100
1000
10000
100000
cumulative family frequency