Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Memory and Language 53 (2005) 496–512

Journal of Memory and Language

www.elsevier.com/locate/jml

Changing places: A cross-language perspective on frequency and family size in Dutch and Hebrew

Fermín Moscoso del Prado Martín [a,*], Avital Deutsch [b], Ram Frost [b], Robert Schreuder [c], Nivja H. De Jong [c,d], R. Harald Baayen [c,d]

[a] *MRC-Cognition and Brain Sciences Unit, Cambridge, UK*
[b] *The Hebrew University, Jerusalem, Israel*
[c] *University of Nijmegen, The Netherlands*
[d] *Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands*

Received 7 August 2003; revision received 20 June 2005
Available online 30 September 2005

Abstract

This study uses the morphological family size effect as a tool for exploring the degree of isomorphism in the networks of morphologically related words in the Hebrew and Dutch mental lexicon. Hebrew and Dutch are genetically unrelated, and they structure their morphologically complex words in very different ways. Two visual lexical decision experiments document substantial cross-language predictivity for the family size measure after partialing out the effect of word frequency and word length. Our data show that the morphological family size effect is not restricted to Indo-European languages but extends to languages with non-concatenative morphology. In Hebrew, a new inhibitory component of the family size effect emerged that arises when a Hebrew root participates in different semantic fields.

Keywords: Morphological processing; Family size effect; Hebrew; Semitic root; Visual lexical decision

The morphological family size of a word is the type count of all the complex words in which this word appears as a constituent (Baayen, Lieber, & Schreuder, 1997; Schreuder & Baayen, 1997). Words with large morphological families elicit faster response latencies in visual lexical decision than words with small morphological families. For instance, the English word *man* appears as a constituent in nearly 200 other English words, including compounds such as *policeman* or *manpower*, and derived words such as *manhood*. On the other extreme, word such as *scythe* do not appear as constituents in any other words. (Note here that inflectional variants of a word are not counted as different members of the morphological family, thus one would not count *scythes* or *policemen* as different from *scythe* or *policeman* for the purpose of calculating morphological family sizes.) Although morphological family size is highly correlated with word frequency, i.e., the more frequent a word is, the larger its morphological family size tends to be, the effect of family size is present even after having controlled for the effects of word frequency (Schreuder & Baayen, 1997). The growing body of experimental evidence on the morphological family size effect shows that

* Corresponding author. Fax: +44 1223 359 062.
E-mail address: fermin.moscoso-del-prado-martin@mrc-cbu.cam.ac.uk (F. Moscoso del Prado Martín).

0749-596X/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2005.07.003

the locus of this effect is at the semantic level. Schreuder and Baayen (1997) and Bertram, Schreuder, and Baayen (2000) pointed out, using visual lexical decision, that the exclusion of opaque family members from the family counts improves the correlations with response latencies. Similarly, in priming experiments the morphological family size of a target word interacts with the amount of priming elicited by semantically related prime–target pairs (Feldman & Pastizzo, 2003). De Jong, Schreuder, and Baayen (2000), again using visual lexical decision, provided further evidence that the effect of family size is not mediated by surface form. For instance, the Dutch irregular past participle "ge**vocht**en" (i.e., *fought*) activates only the family of "**vecht**en" (i.e., *to fight*) and not that of "**vocht**" (*moisture*). In addition, homonyms show differential family size effects when presented in a disambiguating context (De Jong, 2002). The effect of family size has been observed not only for Dutch, but also for English (Baayen, Lieber et al., 1997; see also De Jong, Feldman, Schreuder, Pastizzo, & Baayen, 2002) and German (Lüdeling & De Jong, 2002). Additionally, Feldman and Siok (1997, 1999) report a similar effect in visual lexical decision in Chinese, where characters with semantic radicals (a pseudo-morphological unit without a phonetic realization) that are present in many other characters are recognised faster.

Although previous research indicates that the locus of the family size effect is at the semantic and not at the formal level, it remains an open question how general this finding is. It might be the case that, in other languages with different and more complex word formation systems than English or Dutch, the role of the form-related aspects of a complex word might be more important than its semantic characteristics, compared to English or Dutch. An example of such a language is Hebrew, which belongs to the Hamo-Semitic language family, in contrast with the Indoeuropean languages that were previously studied. Previous research in masked priming has shown that morphologically related Hebrew words showed priming independently of their semantic relatedness. This suggests a stronger role of the purely formal factors of Hebrew morphological structure than in other languages.

All Hebrew verbs and most Hebrew nouns and adjectives are comprised of two basic derivational morphemes: the root and the word pattern. The root usually consists of three consonants, while the word pattern consists of a cluster of vowels, or vowels and consonants. The root generally conveys the core meaning of the word, while the word-pattern creates variations on its meaning, and determines the grammatical properties of the word. This combination of root and word patterns is an example of a non-concatenative morphology (Berman, 1978; McCarthy, 1981). The morphological units of Hebrew words cannot generally be found by splitting words into sequences of morphemes as beads on a string. More complex interposition processes are required.

Although word patterns shape the meaning of words, the exact meaning of a word cannot be unequivocally predicted from its constituent morphemes, the root and the word pattern. This is because of the linguistic characteristics of the word patterns on the one hand, and the existence of homonymic roots on the other hand. As to the word patterns, although they determine the core meaning of words, their semantic characteristics, most markedly in the nominal system, can be ambiguous: many of the nominal word patterns can denote more than one semantic category (e.g., the nominal word pattern *-a-e-et* can denote both profession and disease) on the one hand, and a specific semantic category can be expressed by more than one nominal pattern (e.g., profession can be denoted by the nominal patterns *-a-e-et* and *-a–a-*). (The long dash between the two vowels represents a gemination or 'doubling' of the corresponding root consonant.)

An example of such a homonymic root is *G-D-R* (root consonants are shown in upper case throughout this paper), which is present in the words *GaDeR* (*fence*) and *haGDaRa* (*definition*). In contrast, as an example of a root with a constant semantic contribution, consider the root *SH-M-N*, which always carries the core meaning *fat*. This root can be embedded in the nominal pattern *-e-e-* to form the noun *SHeMeN* (*oil*), in the adjectival pattern *-a-e-* to form the adjective *SHaMeN* (*fat*), or in the verbal pattern *hi--i-* creating the verb *hiSHMiN* (*to become fat*).

Previous studies in Hebrew using masked and cross modal priming paradigms have demonstrated the role of the root morpheme in the lexical access of Hebrew complex words, suggesting that root morphemes constitute lexical units in the Hebrew lexicon (Deutsch, Frost, Pollatsek, & Rayner, 2000; Frost, Deutsch, & Forster, 2000; Frost, Forster, & Deutsch, 1997). Furthermore, the findings in forward masked priming experiments indicate that the lexical status of the root morphemes and their role in mediating lexical access is not conditioned by semantic transparency (Frost et al., 1997), although it can interact with semantic effects (Frost, Deutsch, Gilboa, Tannenbaum, & Marslen-Wilson, 2000). In addition, further investigation of irregular root morphemes, the so-called defective roots in which one or two consonants may be absent, indicates that the role of the root in mediating lexical access is very sensitive to the abstract structural characteristics of the morpheme (Frost et al., 2000). Violation of the standard morphological form inhibits the priming effect between morphologically related words. Possibly, the exact repetition of the root structure allows the root to acquire a formal representation that is independent of meaning. In other words, the Hebrew root seems to be an autonomous form unit, in the sense of Aronoff (1994).

If the representation of the root is indeed independent of meaning, meaning overlap would not be a requirement for a word to participate in a given morphological family. We will refer to this possibility as the *formal family hypothesis*. On the other hand, considering that Hebrew words sharing a root generally belong to the same semantic field, and considering that the morphological priming effect in Hebrew interacts with semantic transparency (Frost, Deutsch, Gilboa et al., 2000), it is also possible that the family size effect in Hebrew might have a conceptual, post-access origin, just as in Germanic languages. According to this second hypothesis, that we will call the *conceptual family hypothesis*, Hebrew roots which express more than one semantic field should be split into as many homonymic roots as there are semantic fields, similar to previous findings in Indoeuropean languages.

In Experiment 1, using visual lexical decision, we clarify the nature of the family size effect in Hebrew, i.e., whether it is formal or semantic in nature. We do this by contrasting non-homonymic roots (i.e., roots that appear in words that always share the same semantic field), with homonymic roots (i.e., roots that appear in words belonging to more than one semantic field). If the conceptual family hypothesis is correct, we expect to find differential family size effects for these two kinds of roots. Non-homonymic roots should reveal a family size effect just as observed in visual lexical decision for Dutch, German, and English. Homonymic roots, on the other hand, might reveal separate family size effects for each of their semantic fields (cf., De Jong, 2002). Alternatively, if the formal family hypothesis is correct, it makes no sense to distinguish between homonymic and non-homonymic roots in Hebrew. In that case, the family size effect will be caused by the total number of words derived from the homonymic root (regardless of the different semantic fields) just as in the case of non-homonymic roots.

We have opted for using visual lexical decision as our experimental task because in our previous studies visual lexical decision has emerged as a highly sensitive task for tracing the effects of family size. Effects of family size and family frequency have also been studied using MEG, see Pylkkänen, Feintuch, Hopkins, and Marantz (2004) and word naming (Baayen, 2005).

The second issue that this study addresses is the extent to which Hebrew and Dutch are isomorphic at the conceptual level. The issue of cross language isomorphy at the conceptual level is addressed by Bates et al. (2003). They report that in picture naming in seven different languages, naming latencies correlate with the (objective or subjective) frequency of a word in a language equally strongly as they correlate with the frequency of its translation into another language. They found, for instance, that Chinese frequencies predict English naming latencies just as well as English frequencies do, and vice versa. As noted by Bates et al., although these results suggest that there is a substantial conceptual component to the word frequency effect, this cross-linguistic evidence does not rule out the possibility of form playing a role as well. To what extent perceptual familiarity with the word form on the one hand (Bradley & Forster, 1987; Morton, 1969) and familiarity with specific objects and concepts on the other hand (Becker, 1979; Borowsky & Besner, 1993; Plaut & Booth, 2000; Stanovich & West, 1981) contribute to this symmetric word frequency effect in picture naming across languages is at present unclear.

To gain further insight into the extent to which the conceptual systems of unrelated languages might be isomorphic, we use the family size effect as a diagnostic tool. Parallel to Experiment 1, we report a second experiment with the Dutch translations of the Hebrew words. This experiment will allow us to ascertain whether the cross-linguistic predictivity of word frequency in picture naming reported by Bates and colleagues generalizes to visual lexical decision. More importantly, this experiment will also make it possible to investigate the cross-linguistic predictivity of another count not considered by Bates and her colleagues, the family size effect. Given the conceptual nature of the family size effect, if Hebrew family sizes predict Dutch response latencies after having partialled out the effects of Dutch word frequency, word length and family size, and, likewise Dutch counts predict Hebrew response latencies, this would provide clear evidence for considerable overlap in cross-linguistic lexico-semantic organization.

For what follows, it should be kept in mind that word formation in Hebrew and Dutch are fundamentally different at the level of form. This difference in form does not only concern the non-concatenative morphology involved in the formation of derived words in Hebrew and the concatenative morphology characterizing word formation in Dutch. A difference between the two languages is also present with respect to the use of compounding for word formation. Compounding is a concatenative process, which is used abundantly in Dutch. In fact, in Dutch, compounding is the most productive word formation process. Although Hebrew has some concatenative derivational affixes, typically used in combination with non-concatenative processes, it makes hardly any use of the purely concatenative process of compounding. The function of compounding as known in English or Dutch is taken over in Hebrew by phrasal constructions. Consequently, morphological families in Hebrew and Dutch differ substantially with respect to both their magnitude and the kind of words that make up the families. In Dutch, as in English, morphological families consist largely of compounds (see Schreuder & Baayen, 1997), while in Hebrew, they consist almost exclusively of derived words. The difference in the magnitude of the family sizes in the two languages is clear from the differences in the ranges. For Dutch, the range of the overall family size counts is [0,549], for Hebrew, the range is roughly [0,25]. If we consider the range of the family size count in Dutch when this count is restricted to derivation, we find a substantially reduced range, [0,11]. This range is smaller than the range for Hebrew, reflecting the greater productivity of derivation in Hebrew compared to Dutch. The bulk of morphological families in Dutch consists of words with compound structure, with a range of [0,448]. In other words, Hebrew and Dutch differ both with respect to the dimension of concatenation and with respect to the productivity of derivation and compounding. Note that, given the smaller sizes of morphological families in Hebrew, it is not self-evident that it will be possible to observe a significant effect in Hebrew. Due to the small family sizes in Hebrew, this lexical variable might not have measurable consequences in this language.

Experiment 1

Method

Participants

Forty undergraduate students at the Hebrew University were paid to take part in this experiment. All were native speakers of Hebrew.

Materials

We compiled a list of 99 Hebrew roots and their morphological families using a Hebrew dictionary (Schweika, 1997). Forty-three of these roots were non-homonymic in the sense that all their family members belong to the same semantic field. The remaining fifty-six roots were homonymic, i.e., their family members belong to two or more different semantic fields. An example of a non-homonymic root is *B-G-R*, which has as family members the words *BaGRut* (*maturity*), *mBuGaR* (*an adult*), *hitBaGRut* (*maturation*), *mitBaGeR* (*teenager*), *BoGeR* (*mature*), *BaGaR* (*to grow up*), *BiGGeR* (*to grow*), and *hitBaGeR* (*to mature*). By homonymic root we refer to Hebrew families whose members cluster into different semantic fields. An example of a homonymic family is the root *X-SH-B*, which appears in two semantic fields, one relating to thinking, e.g., *XaSHaB* (*to think*), *maXSHaBa* (*a thought*), *XaSHiBa* (*thinking*), and one relating to arithmetics and calculations, e.g., *XiSHeB* (*to calculate*), *XeSHBon* (*arithmetics*), *XiSHuB* (*calculation*). It can be observed in this example that, although thinking and calculating both refer to mental activities, any two words from these two semantic fields will be less similar in meaning than any two words selected from the same semantic field.

The assignment of words to semantic fields, carried out by the second author on the basis of dictionary definitions, is supported by an analysis of semantic distances based upon Latent Semantic Analysis (LSA; Landauer & Dumais, 1997). As no operational LSA system is available for Hebrew, we obtained vector-based semantic similarity scores for the English translations of our Hebrew words, using the scripts available at http://lsa.colorado.edu with their default settings. The average similarity between the pairs of target words from different semantic fields of the same Hebrew root was 0.09. We compared these similarities with pairs of words from the same semantic field of a Hebrew root, choosing the target word and a member of the morphological family with, if possible, a translation equivalent in English that was not morphologically related, to be conservative (e.g., *calculation* and *arithmetic*). The average similarity between these pairs of words was 0.31, which is significantly greater than the average similarity score for the pairs of words from different semantic fields ($W = 448.5$, $p < .0001$, one-tailed Wilcoxon rank sum test). In addition, we calculated the LSA similarity scores for the English translations of 25 randomly selected pairs of words from our experiment (independently of them sharing a root or not). The average similarity score between these pairs was also 0.09, not significantly different from the average similarity scores obtained for words from different semantic fields of the same root ($W = 834.5$, $p = .2202$). Thus, our traditional lexicographic assignment of words to semantic fields is supported by current co-occurrence based techniques.

For each non-homonymic root, we selected one family member for presentation in the experiment. For each homonymic root, we selected two words belonging to different semantic fields for inclusion in the experiment. Each of the target words was associated with two counts, the count of family members in the same semantic field, henceforth the *related family size*, and the count of family members in other semantic fields, henceforth the *unrelated family size*.[1] We have chosen this terminology for notational convenience, and it should be kept in mind that there are gradations of semantic similarity both within and between semantic fields. For each word, we also obtained a frequency count using a corpus of 200 million words of Hebrew newspaper texts (David Plaut, personal communication).

We constructed two experimental master lists. Both master lists contained the 43 words representing the 43 non-homonymic roots. To each master list, we added 56 words with homonymic roots. We assigned the words with homonymic roots to these lists such that two words sharing a given homonymic root did not appear in the same list. Thus, exactly the same 99 roots appeared in both lists. We also made sure that the average related family size was approximately the same across the two

[1] Note that, for the unrelated family size count, all non-homonymic roots have a value of zero.

Table 1
Means and standard derivation for the different counts in the Hebrew data set, and in the subsets of homonymic and non-homonymic (after removing the 3 outliers and 1 word with unreliable family size counts)

	Total (151 items)		Non-homonymic roots (43 items)		Homonymic roots (108 items)	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Frequency	32.17	88.44	36.09	111.14	30.63	78.32
Word length	4.16	0.89	3.90	0.91	4.26	0.87
Total family size	12.06	5.94	8.95	5.44	13.28	5.69
Related family size	7.21	4.52	8.95	5.44	6.53	3.93
Unrelated family size	4.79	4.79	0.00	0.00	6.75	4.37
Family size ratio	0.60	1.25	2.14	0.61	0.00	0.86
Response latency	681 ms	56 ms	678 ms	47 ms	682 ms	59 ms
Error	3.02%	4.74%	3.93%	5.22%	2.66%	4.52%

The word frequency counts are in occurrences per million.

lists. Each word was paired with a pseudo-word that did not violate the phonological structure of Hebrew. These pseudo-words were constructed by applying the word-pattern of the corresponding word to a non-existing (but possible) Hebrew root. We constructed three pseudo-randomized versions of each master list, each of which was preceded by a practice session of 10 words and 10 pseudo-words. Both words and pseudo-words were displayed with the vowel signs to avoid ambiguity of interpretation of the target words.Table 1 provides the means and standard deviations for the frequency and family size counts for this dataset.

Procedure

Participants performed the experiment in noise-attenuated experimental rooms. They were asked to decide as quickly and accurately as possible whether the letter string appearing on the computer screen was a real Hebrew word. Each stimulus was preceded by a fixation mark in the middle of the screen for 500 ms. followed after 50ms. blank screen by the stimulus, which appeared at the middle of the screen.

Stimuli were presented on color monitors in white lowercase 28 point letters on a dark background and they remained on the screen for 1500 ms. The maximum time span allowed for a response was 2000 ms from stimulus onset.

Results and discussion

All participants in this experiment performed with an error rate less than 15%. Three words elicited error rates above 30% and were removed from the data set.

Family size is known to be correlated with frequency (Schreuder & Baayen, 1997). In this data set, the correlation between related log family size and log word frequency is $r = .24$ ($p = .0031$). The correlations between word frequency, family size, and word length introduce medium collinearity in our data matrix (with a condition number κ of 15.40), consequently, our analyses are to be interpreted with caution when generalizing beyond the range of frequencies and family sizes represented in the data set (Belsley, Kuh, & Welsch, 1980). To ascertain whether family size contributes to the reaction times independently of word frequency and word length, we first partial out the effect of word frequency and word length before testing for an effect of family size.

For the words with homonymic roots, we have two family size counts, one count for the family members with meanings related to the meaning of the target word (the related family size), and one count of family members with meanings that are unrelated to that of the target (the unrelated family size).

For a by-participant regression analysis, Lorch and Myers (1990) provide a technique (used, e.g., by Alegre & Gordon, 1999) that yields significance values for the predictors, but that does not allow a straightforward procedure for partialling out the contribution of word frequency before assessing the contribution of family size. Hence, we have made use of a multi-level extension of the Lorch and Myers approach that overcomes this disadvantage while maintaining the strengths of the original Lorch and Myers technique (Baayen, Tweedie, & Schreuder, 2002; Pinheiro & Bates, 2000). A multi-level linear regression with log response latency as dependent variable and log word frequency, word length, and log family size as predictors revealed a significant facilitatory effect of word frequency ($F(1, 3716) = 267.8$, $p < .0001$), but no significant effect of word length after having partialled out the effect of word frequency ($F(1, 3715) = 2.1$, $p = .1500$). Both the related ($F(1, 3716) = 10.9$, $p < .0010$) and unrelated family size ($F(1, 3716) = 10.5$, $p < .0012$) counts had significant effects on the regression after partialling out the effect of word frequency.

A by-item multiple regression revealed the same pattern of results: a facilitatory effect of word frequency ($F(1, 146) = 86.3$, $p < .0001$), no effect of word length ($F < 1$), a marginally significant facilitatory effect of the related family size ($F(1,145) = 3.6$, $p = .0598$) and a marginally significant inhibitory effect of the unrelated family size count ($F(1,145) = 2.8$, $p = .0941$), after partialling out the effect of word frequency.

In both the by-participant and by-item regressions, the effects of related and unrelated family size had coefficients of opposite signs: negative for the related family size count, and positive for the unrelated family size count. The absolute values of these coefficients were of very similar magnitudes (−0.010 and 0.008 by-participant; −0.012 and 0.009 by-item). We can test whether these two coefficients differ only in sign but not in absolute magnitude by fitting a second model to the data in which related and unrelated family size are represented by a single parameter, the *family size ratio*,

$$\phi = \log\left(\frac{\text{related family size} + 1}{\text{unrelated family size} + 1}\right), \quad (1)$$

where, as before, we add 1 to the counts to avoid having to take the logarithm of zero. (Throughout the text, we use log to refer to the natural logarithm.) Since

$$\log\left(\frac{\text{related family size} + 1}{\text{unrelated family size} + 1}\right) = \log(\text{related family size} + 1) - \log(\text{unrelated family size} + 1), \quad (2)$$

a model with the family size ratio is mathematically equivalent to a model with two coefficients with different sign but the same absolute value. We compare the more complex model with two predictors with the simpler model with only one predictor (the family size ratio) using a likelihood ratio test for the by-participant analysis (Crawley, 2002, p. 696) and the standard F test for the by-item analysis (Chatterjee, Hadi, & Price, 2000, pp. 70–72).

A multi-level linear regression with log response latency as dependent variable and log word frequency and family size ratio as predictors revealed a significant facilitatory effect of word frequency ($F(1,3717) = 267.9$, $p < .0001$) and a significant facilitatory effect of the family size ratio ($F(1,3717) = 21.3$, $p < .0001$). The likelihood ratio test comparing this simpler model with the more complex model revealed that the extra parameter of the more complex model does not lead to a significantly better fit to the data ($L(7,6) = 0.12$, $p = .7234$).

The by-item multiple regression revealed the same pattern of results: a facilitatory effect of word frequency ($F(1,146) = 87.42$, $p < .0001$) and a significant effect of the family size ratio ($F(1,146) = 7.34$, $p = .0076$) after partialling out the effect of word frequency. The F test comparing the more complex with the simpler model yielded $F < 1$. Again, model parsimony requires selection of the simpler model.

The use of the family size ratio has as additional advantages (1) that its distribution is less skewed—for non-homonymic roots, the unrelated family size count is always zero, and (2) that the collinearity in the model is reduced. Hence, the models with the family size ratio combine greater parsimony with a data matrix that is more appropriate for regression analysis.

The error analysis for this dataset confirmed the main effects that were found on the reaction times, without showing any speed-accuracy trade-off.

The above analyses showed solid effects of word frequency in both the by-subject and the by-item analyses. None of these analyses revealed an independent contribution of Hebrew word length after having partialled out the effects of word frequency. The effect of written frequency, documented in Hebrew in the present paper for the first time, is not surprising given the strong effects of frequency that have been reported in many other languages, and in cognition in general (e.g., Hasher & Zacks, 1984; Taft, 1979). The lack of reliable effects for word length is probably due to the relatively low variance of Hebrew word length in our data.

More interesting is that, in these analyses, the family size ratio also emerged as a strong independent predictor of response latencies, after partialling out the effects of word frequency and word length. In other words, our analyses show that related and unrelated family size have effects of the same magnitude but with opposite sign, facilitatory for the related family, and inhibitory for the unrelated family [see Eq. (1)]. The latter inhibitory effect is, by definition, present only for words with homonymic roots—for non-homonymic roots, the count of unrelated family size is zero, and does therefore not contribute to the family size ratio.

The effect of the family size ratio is consistent with two different interpretations. One possible interpretation is that we are observing the outcome of two separate processes of equal magnitude. Another possibility is that we are observing the effect of a single process, for instance, the joint influence of all words sharing a root, moderated by their semantic relation to the target word. The present data do not enable us to choose between these options, although Occam's razor favors the single process solution. Further research is required here.

Experiment 1 documents for the first time a Family Size effect for a non Indo-European language. Whereas the Family Size effect in Indo-European languages is anchored in the presence of a shared stem, i.e., an independent morphological unit that, in English, German, or Dutch, can even be a word in the language, the family size effect in Hebrew is anchored in the Semitic root, a non-concatenating morphological unit, which can never exist as a word without a word pattern. In addition, Experiment 1 shows that family size effects are present in a language with hardly any compounding. This shows that it is not the formal aspects of word formation (concatenative or non-concatenative, derivation or compounding) that are relevant for the family size effect, but rather the semantic aspect of word formation: the creation or denotation of concepts. Finally, the results for the homonymic roots support the conceptual family hypothesis and argue against the formal family hypothesis.

The inhibitory effect of the unrelated family size that emerges for the Hebrew words with homonymic roots contrasts with the family size effects documented for homonymic words in Dutch (De Jong, 2002). When Dutch homonymic words are presented in a disambiguating context, only the family size of the contextually appropriate meaning has a facilitatory effect, with no trace of an inhibitory effect of the family size of the contextually inappropriate reading of the homonym. Note that there is a difference between the Dutch and Hebrew homonyms. Dutch homonyms, like English homonyms such as *bank*, share the same phonological form, while Hebrew words with homonymic roots have different phonological forms (compare, for instance, *GaDeR* and *haGDaRa*). This suggests that the inhibition from the unrelated family size observed for Hebrew homonymic roots might arise due to the combination of differences in phonological form and differences in meaning for words sharing the same basic morpheme. This hypothesis receives support from a study addressing the family size effect in the bilingual lexicon. Dijkstra, Moscoso del Prado Martín, Schulpen, Schreuder, and Baayen (2005) observed an inhibitory effect in the bilingual lexicon similar to the one reported in the present study for Hebrew words with homonymic roots. Dutch-English interlingual homographs, i.e., words with identical spelling but different meanings and pronunciations across two languages, elicited a facilitatory effect of the homograph's family size in the language that is relevant for the task (e.g., Dutch, when performing Dutch visual lexical decision), and an inhibitory effect of the family size in the irrelevant language (e.g., English, when performing Dutch lexical decision). Crucially, the amount of inhibition caused by the family size in the irrelevant language increases as the pronunciations of the homograph in Dutch and English become more dissimilar. This suggests that the accumulation of dissimilarities at the phonological and semantic levels interferes with lexical decision.

Experiment 2 examined the visual lexical decision latencies to the Dutch translation equivalents of the Hebrew words used in Experiment 1. This second experiment was motivated by three considerations. First, it follows from the claim that the family size effect is semantic in nature that the family size effect observed for Hebrew should also be observed for the translation equivalents of these Hebrew words in Dutch. More specifically, the family size counts in Hebrew might be significant predictors for Dutch response latencies after partialling out Dutch word frequency. If so, this would provide evidence for cross-language similarity in lexical conceptual organization. Second, since the Hebrew words with homonymic roots have morphological families that translate into Dutch words that are not morphologically related, we predict that the Hebrew unrelated family size should not be predictive for the translation equivalents of the Hebrew homonymic words, while the Hebrew related family size should have significant cross-language predictive value. Third, Experiment 2 will enable us to address the issues raised by Bates and her colleagues, namely, whether the frequency counts from a given language predict the response latencies in another unrelated language equally well as the frequency counts from that language itself.

Experiment 2

Method

Participants

Thirty-six undergraduate students at the University of Nijmegen were paid to take part in this experiment. All were native speakers of Dutch.

Materials

The 155 Hebrew words from Experiment 1 were translated into Dutch. The translation was done by hand using a Dutch-Hebrew Dictionary (Bolle & Pimentel, 1984), checked with the English translation of the words provided by the second author, and validated by a Hebrew-Dutch bilingual. When a word had different possible translations into Dutch with different meanings, we included all different translations in the experiment. In this way, after removing those words which appeared twice, we obtained a set of 162 Dutch words. Frequency and family size counts for these words were extracted from the CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1995). (The family size count of a given word is calculated by counting the number of morphological parses in the CELEX lexical database in which the stem appears as a constituent, independently of their regularity or semantic relation to the stem.) Each of these words was paired with a pseudo-word whose phonotactics did not violate the phonology of Dutch. These pseudo-words were constructed by changing one or two letters of the stem of the corresponding Dutch word without altering its affixes. Twenty practice trials, ten words and ten pseudo-words were run before the actual experiment. We constructed three different permutations and their corresponding reversed versions of the original word list.

Table 2 provides the means and standard deviations for word length, logarithmic frequency, and logarithmic family size counts for this data set.

Table 2
Means and standard derivation for the different counts after removing the three outliers in the Dutch data set

160 items	Mean	Standard deviation
Frequency	88.26	312.55
Word length	7.17	2.47
Family size	24.81	30.56
Response latency	578 ms	78 ms
Error	3.17%	4.71%

Procedure

Participants performed the experiment in noise-attenuated experimental rooms. They were asked to decide as quickly and accurately as possible whether the letter string appearing on the computer screen was a real Dutch word. Following a pause after the test trials, the experiment was run with two further pauses, dividing the experiment in three blocks, each of them containing one third of the stimuli. Each stimulus was preceded by a fixation mark in the middle of the screen for 500 ms. After these 500 ms., the stimulus appeared at the same position. Stimuli were presented on NEC Multisync color monitors in white lowercase 21 point Arial letters on a dark background and they remained on the screen for 1500 ms. The maximum time span allowed for a response was 2000 ms from stimulus onset.

Results and discussion

All participants in this experiment performed with an error rate less than 15%. Three words elicited error rates above 30% and were removed from the data set.

A by-participant multilevel linear regression analysis of the data with log reaction time as dependent variable and log frequency, word length, and log family size as independent covariates revealed facilitatory main effects of word frequency ($F(1, 5505) = 773.18$, $p < .0001$), an inhibitory effect of word length ($F(1, 5505) = 247.98$, $p < .0001$, after partialling out the effect of word frequency), and a facilitatory effect of family size ($F(1, 5505) = 62.413$, $p < .0001$, after partialling out the effects of word frequency and word length).

A by-item multiple regression with log reaction times as the dependent variable and logarithmic word frequency, word length, and logarithmic family size similarly showed a facilitatory main effect of word frequency ($F(1, 156) = 138.451$, $p < .0001$), an inhibitory main effect of word length ($F(1, 156) = 42.82$, $p < .0001$, after the effect of word frequency), and a facilitatory effect of family size ($F(1, 156) = 10.97$, $p = .0011$, after partialling out the effects of word frequency and word length).

The error analysis for this dataset confirmed the main effects that were found on the reaction times, without showing any speed-accuracy trade-off.

We further inspected the contributions of the (small) counts of derived words in the morphological families and the (large) counts of the compounds in the morphological families separately. A multiple regression model on the item means revealed a small effect in the expected direction for the derived family size ($F(1, 154) = 2.85) = 0.0468$, one-tailed test and a highly significant effect for the compound family size ($F(1, 154) = 15.71$, $p < .0001$, two-tailed test) in the expected direction. This result shows that in Dutch the family size effect is not restricted to compounds only. Even for the small numbers of derived words, an effect of family size seems to be present.

Summing up, just as for Hebrew, the analysis of the Dutch data set revealed strong facilitatory effects of written word frequency and family size, i.e., frequent words and words with large morphological families elicit shorter response latencies. Additionally, in Dutch, we also found an inhibitory effect of word length, i.e., longer words elicit longer response latencies.

Cross-language analyses

We now consider the similarities and differences between family size, word length, and frequency effects in Hebrew and Dutch. Recall that Bates and her colleagues report that picture naming latencies in several unrelated languages can be predicted equally well from frequency counts from the same language as from frequency counts of translation equivalents in another (unrelated) language. The question to be addressed now is whether a similar symmetry in cross-language predictivity is present in our comprehension data for Hebrew and Dutch.

In what follows, we first study the correlations between our independent variables in both experiments. This will allow us to gauge to what extent the frequency, family size, and word length counts in the two languages interrelate. We continue by assessing the individual predictive power of each of these variables on the response latencies in both experiments. Finally, we present multiple regression analyses in which we assess whether each of these variables has additional predictive power on the response latencies, after having partialled out the effects of the within language counts that were found in the within language analyses.

Before reporting these analysis, two comments on the variables in our study are in order. First, recall that the Hebrew and Dutch frequency counts are based on corpora of different size, 200 million words for Hebrew and 42 million words for Dutch. A preliminary question is to what extent this difference might affect our cross-linguistic comparisons. With respect to the frequency counts, Baayen, Moscoso del Prado, Schreuder, and Wurm (2003) show that under the lognormal model (Baayen, 2001; Carroll, 1967) a change in corpus size amounts to a change in the intercept for the regression of frequencies in the one language on the frequencies in the other language. Since our analyses make use of log-transformed frequencies, the change in corpus size has no qualitative effect on our regression models. However, there will be more sampling error in the Dutch counts (which are based on the smaller corpus), but this may be offset by better morphological analyses and better corpus design for Dutch than for Hebrew, resulting in more accurate frequency estimates.

With respect to the family size counts, it should be kept in mind that these are based on a comparative analysis of several dictionaries for Hebrew, and on the morphological parses available for Dutch in the CELEX lexical database. For both languages, we have used the most comprehensive resources available for estimating the family size counts. Note that the typological difference between Hebrew and Dutch is visible in the family size counts, which in Hebrew are restricted to a smaller range ([0,3.37], in log units) than in Dutch ([0,4.82], in log units). Under the simplifying assumption that a one unit change in family size has the same effect size in both languages, the Dutch family size has a larger a-priori probability of being a significant predictor. The same holds for length in letters (range [3,18] for Dutch, range [3,7] for Hebrew). Therefore, observing cross-language effects for the Hebrew variables, especially family size and word length, is potentially more difficult than for the Dutch variables.

We begin with an analysis of the correlational structure of the Hebrew and Dutch frequency, word length, and family size measures. Table 3 summarizes the correlations between the logarithmic frequency counts, word length, and logarithmic family size in Hebrew and Dutch. We find an asymmetry in the correlations between Dutch family size and the word frequencies in both languages. On the one hand, according to Downie and Heath's (1965) method for comparing correlation coefficients from non-independent samples, the correlation of Dutch family size and Dutch frequency ($r = .63$) is significantly stronger ($Z = 2.88$, $p = .0020$) than the correlation between Dutch family size and Hebrew frequency ($r = .40$). On the other hand, the correlations of Dutch or Hebrew frequency with Hebrew related family size do not differ significantly ($Z = 0.34$, $p = .3657$). In fact, the correlation between Dutch family size and Hebrew frequency ($r = .40$) is significantly stronger ($Z = 1.72$, $p = .0428$) than the correlation between Hebrew related family size and Hebrew frequency ($r = .24$).

With respect to word length, we observed a clear pattern. Dutch word length correlates better than Hebrew word length both with Hebrew frequency ($Z = 1.74$, $p = .0409$) and with Dutch Frequency ($Z = 2.81$, $p = .0024$), with the correlation between Hebrew length and Hebrew frequency not reaching significance ($r = -.14$, $p = .08$).

Table 4 summarizes the Pearson correlations between the frequency and family size counts and the response latencies in the two languages. First consider the correlations between the frequency measures and the reaction times in both languages. Note that there is an asymmetry between the correlation of frequency and response latencies. The correlations of Hebrew frequency and Hebrew RT ($r = -.60$) and Dutch frequency and Dutch RT ($r = -.65$) are not different according to Fisher's Z-transformation for comparing correlation coefficients ($Z = 0.62$, $p = .2593$). By contrast, Dutch frequency shows a correlation with Hebrew RT ($r = -.39$) that is significantly smaller ($Z = 2.57$, $p = .0051$) than the correlation of Hebrew frequency and Hebrew RT ($r = -.60$). Similarly the correlation between Hebrew frequency and Dutch RT ($r = -.44$) is significantly reduced ($Z = 2.61$, $p = .0051$) when

Table 4
Pairwise Pearson correlation coefficients between the frequency, word length, and family size measures in both languages, with mean logarithmic response latencies from both experiments

	RT Hebrew	RT Dutch
Frequency (Hebrew)	−0.60*	−0.44*
Frequency (Dutch)	−0.39*	−0.65*
Related family size (Hebrew)	−0.26*	−0.26*
Family size ratio (Hebrew)	−0.21*	−0.14
Family size (Dutch)	−0.36*	−0.60*
Word length (Hebrew)	0.14	0.17*
Word Length (Dutch)	0.30*	0.60*

The correlations marked with an asterisk are those that are significant at the 5% level.

Table 3
Pairwise Pearson correlation coefficients between the logarithmic frequency, word length, and family size counts

	Frequency Hebrew	Frequency Dutch	Length Hebrew	Length Dutch	Fam. S. Hebrew	Fam. S. Ratio Hebrew
Fam. S. Dutch	0.40*	0.63*	−0.15	−0.50*	0.33*	0.17*
Fam. S. Ratio Hebrew	0.06	0.07	−0.04	−0.02	0.04	
Rel. Fam. S. Hebrew	0.24*	0.21*	0.11	−0.11		
Length Dutch	−0.31*	−0.44*	0.30*			
Length Hebrew	−0.14	−0.18*				
Freq. Dutch	0.46*					

The correlations marked with an asterisk are those that are significant at the 5% level.

compared with the correlation between Dutch frequency and Dutch response latencies ($r = -.65$). In other words, correlations between frequency and reaction times are weaker across languages than within languages. This pattern differs from that reported by Bates et al. (2003) for the picture naming paradigm: they observed equal cross and within-language correlations.

A pattern different from that observed for the frequency-RT correlations emerges in the correlations between family sizes and reaction times, as can be seen in the central sector of Table 4. The correlation of Dutch family size with Dutch response latencies ($r = -.60$) is significantly stronger ($Z = 4.05$, $p < .0001$) than that of Hebrew family size ratio with Hebrew response latencies ($r = -.21$), (this is also true for Hebrew related family size and Hebrew reaction times ($r = -.26$, $Z = 3.55$, $p = .0002$) or Hebrew total family size ($r = -.16$, $Z = 4.42$, $p < .0001$)). At the same time, the correlation between Dutch family size and Hebrew response latencies (-0.36) is slightly greater than the correlation between Hebrew family size ratios and Hebrew response latencies (-0.21). Although this difference is only marginally significant ($Z = 1.55$, $p = .06$), it is interesting since it indicates that Dutch family size is at least as good a predictor of Hebrew response latencies, as any of the Hebrew family size counts.

Word length (see the bottom rows of Table 4) shows a similar pattern to that of family size. The correlation between Dutch word length with Dutch reaction times ($r = .60$) is significantly greater ($Z = 4.65$, $p < .0001$) than that between Hebrew word length and Hebrew reaction times ($r = .14$). At the same time, the correlation between Dutch word length and Hebrew reaction times ($r = .30$) is greater ($Z = 1.52$, $p < .05$) than the correlation between Hebrew word length and Hebrew reaction times ($r = .14$). This indicates that Dutch word length emerges as a better predictor of response latencies in both languages. The reduced correlations observed for Hebrew family size and Hebrew length are in line with the reduced ranges of these variables in Hebrew as compared to Dutch.

Having outlined the correlational structure of the data sets, we proceed to ascertain whether frequency, word length and family size are significant predictors of the response latencies in the other language. More precisely, we want to ascertain whether frequency, word length, and family size counts from one language, explain variance in the response latencies in the other language, after having partialled out the within language variables. For instance, we may ask whether Dutch frequency, family size and word length predict Hebrew response latencies, after having partialled out Hebrew frequency, word length and family size.

For the by-participant analysis, we added log Hebrew frequency, Hebrew word length, and log Hebrew related family size, one at time, as predictors to the by-participant regression fit to the Dutch data in Experiment 2. Sequential analyses of variance revealed additional significant main effects of Hebrew frequency ($F(1, 5221) = 13.38$, $p = .0003$), and Hebrew related family size ($F(1, 5221) = 20.31$, $p < .0001$), without any significant effect for Hebrew word length ($F < 1$) or family size ratio ($F(1, 5220) = 3.78$, $p = .0518$).

For the by-item analysis, we similarly added log Hebrew word frequency, Hebrew word length, and Hebrew family size ratio (one at a time) as predictors to the by-item regression model that was fit to the Dutch data in Experiment 2. Sequential analyses of variance revealed an additional significant facilitatory effect of Hebrew related family size ($F(1, 145) = 3.97$, $p = .0482$), and did not reveal any additional main effects of log Hebrew word frequency ($F(1, 145) = 2.07$, $p = .1521$), Hebrew word length ($F < 1$) or family size ratio ($F < 1$).

What these analyses show are a strong facilitatory effect of Hebrew related family size on Dutch response latencies and an indication of a similar effect of word frequency. Crucially, we do not find an inhibitory effect of the number of unrelated family members in Hebrew on Dutch reaction times. This is in line with our interpretation of the family size effect in Hebrew as a complex effect with an inhibitory and a facilitating component. The facilitating component is not language specific. Due to the translation equivalence between Hebrew and Dutch target words, sizes of the morphological families in the two languages are correlated, consequently, Hebrew related family size predicts Dutch response latencies. The inhibitory component is specific to Hebrew. As a Hebrew speaker activates all words sharing a given root, the unrelated family members are activated along with the related family members. We have seen that the related family members give rise to facilitation in the response latencies, and that the unrelated family members give rise to inhibition. This inhibitory effect does not arise in Dutch because the unrelated family members of the Hebrew target words fall completely outside the morphological families of their Dutch translation equivalents.

To study the predictive power of Dutch word frequency, word length and family size on Hebrew response latencies in more detail, we added the Dutch family counts and frequencies from Experiment 2 as predictors for the Hebrew response latencies of Experiment 1.

By-participant multilevel regression analyses revealed additional significant main effects of Dutch frequency ($F(1, 3716) = 10.58$, $p = .0012$), Dutch word length ($F(1, 3715) = 9.71$, $p = .0019$), and Dutch family size ($F(1, 3716) = 11.01$, $p = .0009$) in sequential analyses of variance, i.e., after having partialled out the Hebrew predictors.

By-item regression analyses revealed only marginal main effects of log Dutch word frequency ($F(1, 145) =$

3.83, $p = .0523$), Dutch family size ($F(1,145) = 2.83$, $p = .0944$), and Dutch word length ($F(1,145) = 3.53$, $p = .0623$) in sequential analyses of variance, under conservative two-tailed testing (as predicted, the effect of family size is also facilitatory).

These analyses show that frequency is a significant predictor across the two languages. In addition, related family size in Hebrew and family size in Dutch emerge as cross-language predictors. We interpret these two family size measures as capturing cross-language similarities in semantic space. Note that the morphology of Hebrew introduces an asymmetry for the Hebrew homonymic roots, with cross-language predictivity for the related family size but not for the unrelated family size.

General discussion

In this study, we have addressed two main questions. First, is family size in Hebrew determined only by morphological form or also by meaning? Second, do the conceptual structures that give rise to the family size effect overlap across languages?

In Experiment 1, we used a regression design with frequency, word length, and family size as independent variables, and response latencies and error scores as dependent variables. We observed independent effects of frequency and family size, both for the response latencies and for the error measure. The fact that family size emerged as an independent predictor supports the hypothesis advanced in Feldman, Frost, and Pnini (1995) that the number of words in which a Hebrew root appears might be an important processing variable. The predictivity of the family size count for a language with non-concatenative morphology and with derivation to the near exclusion of compounding supports the hypothesis that the family size effect is independent of word form and arises at post-access, central levels of lexical processing and representation.

This hypothesis finds further support by the interesting new finding that the effect of family size was different for the non-homonymic roots and the homonymic roots. For non-homonymic roots, the family comprises, by definition, only semantically related words. The count of this number of words is negatively correlated with reaction times. In the case of homonymic roots, the count of all family members does not correlate well with reaction times. However, when we count the semantically related family members and the semantically unrelated family members separately, both counts correlate with the reaction times and errors, with equivalent strength in absolute value, but with opposite signs. The related family size is facilitatory, while the unrelated family size is inhibitory. These different effects of related and unrelated family size counts argue in favor of the conceptual family hypothesis. According to this hypothesis, the family size effect in Hebrew arises at the conceptual level, just as it does in Dutch or English. However, the presence of an inhibitory effect for the unrelated family members suggests that Hebrew speakers are indeed sensitive to the formal aspect of the three-consonantal roots of Hebrew words. This is partially in line with the formal family hypothesis: it suggests that sharing a purely formal root leads to activation of all family members with that root. However, contrary to the formal family hypothesis, the unrelated family members give rise to longer instead of shorter RT's, indicating that Hebrew speakers are also sensitive to the degree of semantic relatedness between the morphologically related words.

At this point, the question arises of whether the inhibitory effect of the unrelated family size arises at the level of orthographic or phonological form, or at the conceptual level. Consider the possibility that it arises at the form level, keeping in mind that the semantic distinction between related and unrelated family size members cannot be operative at the pre-lexical level of form processing by definition. In this case, one would predict that a greater total family size would lead to greater inhibition, as all family members share the same root form. This prediction, however, holds true only for the unrelated family members, and not for the related family members for words with homonymic and non-homonymic roots alike. This points to competition arising at more central, post-access levels of lexical processing. Since the formal aspects of the Hebrew root play a central role in the process of word recognition, we hypothesize that the activation of a homonymic root results in activation of conflicting semantic fields and hence to semantic competition (see, e.g., Gaskell & Marslen-Wilson, 1997; Plaut & Booth, 2000, for more detailed theories of competition at the semantic level). Interestingly, this pattern of effects in opposite directions, i.e., facilitatory for the related members of the family and inhibitory for the unrelated family members, is reminiscent of the results reported for Chinese semantic radicals. In priming experiments, Feldman and Siok (1999) found that Chinese characters sharing the same semantic radical with the target character facilitated its recognition when both prime and target where semantically related, but inhibited the recognition of the target if prime and target where semantically unrelated. (but see Taft & Zhu, 1997, for data suggesting that the semantics of the Chinese radical might be irrelevant).

Experiment 2 investigated the processing in the visual modality of the Dutch translation equivalents of the Hebrew words from Experiment 1. We observed independent effects of written frequency and family size. Interestingly, cross-language comparisons revealed that reaction times for Hebrew words can be predicted from the frequency and family size counts of their Dutch translation equivalents, and vice-versa. Crucially, this is so even after having partialled out the effects of the within language variables. The predictivity of frequency across languages in visual lexical decision is in line with the results reported by Bates et al. (2003) for picture naming.

Recall that Bates and colleagues report equal predictivity for word frequency within and across languages, whereas we observed asymmetrical cross-language predictivity. There are several possible reasons for this difference. In picture naming, participants receive exactly the same pictorial visual input, independently of the language in which the pictures are being named. In visual lexical decision, by contrast, participants receive orthographic visual input which differs substantially between Hebrew and Dutch. Cross-language differences are more likely to arise in a task with language-specific stimuli than in a task using non-linguistic stimuli that are language-independent. More specifically, in picture naming, participants begin with conceptualizing the picture shown, and there is little reason to suppose that this process would differ substantially across speakers of different languages, especially as Bates et al. (2003) used pictures that were explicitly selected by their 'potential cross-cultural validity.' Conversely, in visual lexical decision, participants have to process the visual input, which, in the case of Hebrew, proceeds primarily through root consonants that do not even exist in Dutch. In other words, language differences are present from stimulus onset in visual lexical decision, which may make it more likely to elicit cross-language differences in lexical processing.

Furthermore, a comparison of the concreteness and imageability ratings for the materials used by Bates and the English translations of the materials we used for Hebrew and Dutch points to substantial differences ($t = 7.89$, $df = 43.681$, $p < .0001$ for concreteness, $t = 7.01$, $df = 44.81$, $p < .0001$ for imageability). These differences are not surprising given that Bates et al. used picture naming, which requires narrowing down the stimuli to picturable, concrete, and highly imageable nouns. Our materials were unrestricted in this sense. They varied across word categories, and some words had culturally highly specific meanings in Hebrew, e.g., *orthodox, sermon, parable*, and *garbage site* (*mizbala*), a special place for depositing used bibles. Thus, our cross-linguistic comparison is more susceptible to cultural-specific differences in the usage (and perhaps meaning) of the translation equivalents.

Whereas the cross-linguistic predictivity of word frequency provides only inconclusive evidence for the similarity and salience of concepts across languages, the cross-language predictivity of family size documented in the present study provides unambiguous evidence for a surprising degree of isomorphism in the way concepts are organized in the Hebrew and Dutch mental lexicon. Although these two languages are genetically unrelated and make use of radically different means for creating morphologically complex words (non-concatenative versus concatenative, and derivation versus predominantly compounding), the networks of morphologically related words are similar enough to allow response latencies in the one language to be predicted from the network size in the other language. Thus, the family size effect emerges as an excellent tool for mapping the degree of isomorphism in the conceptual relations in the Hebrew and Dutch mental lexicons.

Acknowledgments

The authors are indebted to Keren Shatzman for her help in translating the Hebrew stimuli into Dutch. We also wish to thank Laurie Feldman, David Plaut, and three anonymous reviewers for comments and suggestions on earlier versions of this manuscript. This research was supported by the Dutch Research Council (NWO), the Max Planck Institute for Psycholinguistics, and the University of Nijmegen, through a PIONIER Grant to R.H. Baayen. During the last stages of preparation of this manuscript, the first author received additional support from the Medical Research Council (UK) and the European Community under the "Information Society Technologies Programme" (IST-2001-35282).

Appendix A

List of stimuli used in the experiments

Dutch			Hebrew					English
Word	Freq.	Fam. size	Root	Word	Freq.	Unrelated	Related	Word
aantal	16,009	0	$s-p-r$	[*mispar*]	88212	23	13	"number"
adoptie	119	5	$\dot{h}-m-\check{s}$	[*ḣimuš*]	1,222	7	5	"adoption"
afdeling	2,795	23	$\dot{h}-g-p$	[*ḣagap*]	6,470	5	2	"wing"

(*continued on next page*)

Appendix A (*continued*)

Dutch			Hebrew					English
Word	Freq.	Fam. size	Root	Word	Freq.	Unrelated	Related	Word
afzetter	7	2	*g − z − l*	[*gazlan*]	2	0	8	"swindler"
angst	7,490	42	*x − r − d*	[*xarada*]	3,313	4	7	"anxiety"
annulering	11	1	*b−T − l*	[*biTul*]	4,169	0	12	"cancellation"
arbeider	2,606	58	*'− b − d*	[*'oved*]	17,424	6	8	"worker"
archief	453	13	*g − n − z*	[*gniza*]	47	0	8	"archives"
baard	974	15	*z − q − n*	[*zaqan*]	3,071	8	3	"beard"
balans	598	9	*ḣ − z − n*	[*maḣazan*]	2,268	5	8	"balance"
barman	106	1	*m − z − g*	[*mozeg*]	75	7	6	"barman"
barst	373	9	*b − q −'*	[*beqa'*]	349	0	16	"crack"
been	7,991	76	*r− g − l*	[*regel*]	7,313	8	9	"leg"
bemesting	25	1	*z − b − l*	[*zibul*]	0	2	5	"fertilizing"
bemesting	25	1	*d − ṡ − n*	[*diṡun*]	355	0	8	"fertilization"
berekening	689	11	*x − ṡ − b*	[*xiṡub*]	1,297	10	15	"calculation"
bevelschrift	9	1	*g − z − r*	[*gzera*]	722	9	4	"edict"
bewerking	374	7	*'− b − d*	[*'ibud*]	3,076	8	6	"processing"
bezoek	4,163	18	*b − q − r*	[*biqur*]	8,018	8	3	"visit"
biggelen	42	2	*z − l − g*	[*zalag*]	1	12	0	"to shed"
brief	8,494	78	*k − t − b*	[*mixtab*]	11,680	0	29	"letter"
brug	2,189	37	*g − ṡ − r*	[*geṡer*]	4,500	0	5	"bridge"
cadet	460	1	*š − ' − r*	[*šo'er*]	58	5	1	"cadet"
chagrijning	109	2	*z − ' − p*	[*zo'ep*]	87	0	13	"enraged"
cirkel	1,115	21	*'− g −l*	[*'igul*]	512	3	12	"circle"
controle	2,307	47	*b − q − r*	[*baqara*]	1,295	4	7	"control"
cultuur	4,200	104	*g − d − l*	[*gidul*]	10,717	16	5	"cultivation"
definitie	1,441	5	*g − d − r*	[*hagdara*]	1,456	8	5	"definition"
dief	572	21	*g − n − b*	[*ganab*]	1,640	0	13	"thief"
ding	15,709	16	*d − b − r*	[*dabar*]	98,082	17	2	"thing"
doof	377	16	*x − r − ṡ*	[*xereṡ*]	343	9	8	"deaf"
dragen	11,963	82	*l − b − ṡ*	[*labaṡ*]	743	0	14	"wore"
dwingen	3,451	6	*ḣ − l − š*	[*ḣileš*]	286	0	6	"to compel"
eenheid	2,984	54	*ḣ − x − d*	[*ḣaxdut*]	22,945	0	12	"unity"
eindigen	2,041	9	*g − m − r*	[*gamar*]	7,209	0	5	"to finish"
eis	3,420	20	*d − r − ṡ*	[*driṡa*]	2,877	8	6	"requirement"
fusie	236	4	*m − z − g*	[*hitmazgut*]	51	10	3	"coalescence"
gaande	416	1	*h − l − k*	[*halika*]	3,563	1	16	"walking"
gebruiken	17,227	53	*ṡ − m − ṡ*	[*hiṡtameṡ*]	1,541	3	12	"used"
gedachte	11,735	23	*x − ṡ − b*	[max*ṡaba*]	4,744	14	11	"thought"
geleidingsvermogen	4	1	*h − l − k*	[*molikut*]	71	14	3	"conductivity"
gevecht	1,025	38	*ḣ − b − q*	[*maḣabaq*]	6,067	10	6	"struggle"
gewoonte	2,459	21	*r − g − l*	[*hergel*]	4,259	9	8	"habit"
gezwel	249	21	*g − d − l*	[*gidul*]	10,717	16	5	"growth"
glanzend	0	14	*b − r − q*	[*mabriq*]	2,608	1	6	"shining"
gooien	4,594	24	*z − r − q*	[*zaraq*]	799	5	6	"to throw"
greep	2,231	14	*ḣ − x − z*	[*ḣaxiza*]	1,791	2	4	"grasp"
grendel	304	7	*b − r − x*	[*bariax*]	664	6	3	"bolt"
groot	51,667	100	*g − d − l*	[*gadol*]	62,129	5	16	"big"
haan	717	14	*g − b − r*	[*geber*]	12,899	12	3	"rooster"
hechtenis	95	2	*x − b − ḣ*	[*maxboṡ*]	19	4	4	"detention"
heilig	2,261	26	*q − d − ṡ*	[*qadoṡ*]	3,051	3	11	"holy"
hek	1,307	12	*g − d − r*	[*gader*]	1,710	7	6	"fence"
held	1,407	29	*g − b − r*	[*gibor*]	4,067	9	6	"hero"
herstelling	15	3	*t − q − n*	[*tiqun*]	4,375	7	11	"repair"
hoffelijk	248	3	*ḣ − d − v*	[*ḣadiv*]	478	0	2	"polite"
hypothese	789	10	*ṡ − ' − r*	[*haṡ'ara*]	863	3	4	"hypothesis"
indruk	6,556	7	*r − ṡ − m*	[*marṡim*]	6,275	12	6	"impression"
ingewikkeld	1,951	1	*s − b − k*	[*mesubak*]	2,659	2	8	"complicated"

Appendix A (*continued*)

Dutch			Hebrew					English
Word	Freq.	Fam. size	Root	Word	Freq.	Unrelated	Related	Word
injectiespuit	62	1	*z − r − q*	[*mazreq*]	417	6	5	“injector”
keuken	4,040	57	*T − b − x*	[*miTbax*]	2,863	4	4	“kitchen”
kledingstuk	377	0	*b − g − d*	[*beged*]	2,391	7	2	“cloth”
knipsel	71	6	*g − z − r*	[*gazir*]	31	6	7	“clippable”
koesteren	1,575	3	*T − p − x*	[*Tipeax*]	134	2	4	“to nurture”
koken	2,436	30	*b − ṡ − l*	[*bišul*]	4,310	2	12	“cooking”
koning	4,249	55	*m − l − k*	[*melek*]	6,764	0	12	“king”
kristallisatie	14	2	*g − b − ṡ*	[*gibuṡ*]	980	0	11	“crystallization”
kus	784	11	*n − ṡ − q*	[*nšiqa*]	1,238	9	7	“kiss”
laatst	11,769	4	*ḣ − x − r*	[*ḣaxaron*]	9,619	6	3	“last”
lampenkap	0	1	*ḣ − h − l*	[*ḣahil*]	155	2	2	“lampshade”
landgoed	407	2	*ḣ − x − z*	[*ḣaxuza*]	342	3	3	“estate”
liefde	7,199	70	*ḣ − h − v*	[*ḣahava*]	21,292	0	10	“love”
lijm	284	10	*d − b − q*	[*debeq*]	2,217	3	14	“glue”
lijst	1,858	38	*s − g − r*	[*misgeret*]	3,884	15	7	“frame”
lijst	1,858	38	*r − ṡ − m*	[*ršima*]	5,168	6	12	“list”
man	50,439	347	*g − b − r*	[*geber*]	12,899	12	3	“man”
minuscuul	165	0	*z − ‘− m*	[*za‘um*]	564	6	1	“tiny”
model	4,370	31	*d − g − m*	[*degem*]	3,106	0	10	“model”
moed	2,101	53	*ḣ − m − š*	[*ḣomeš*]	3,369	1	11	“courage”
moordpartij	59	0	*T − b − x*	[*Tebax*]	2,070	4	4	“slaughter”
neusring	0	0	*n − z − m*	[*nezem*]	17	0	1	“nose-ring”
omtrekken	37	4	*ḣ − g − p*	[*ḣigup*]	30	2	5	“outflaking”
onverschillig	1,174	1	*ḣ − d − ṡ*	[*ḣadiš*]	990	0	2	“indifferent”
onderzoeken	4,017	0	*b − d − k*	[*badak*]	15	10	10	“to examine”
oogst	674	16	*q − š − r*	[*qašir*]	1,420	15	5	“harvest”
oor	4,671	66	*ḣ − z − n*	[*ḣozen*]	1,691	8	5	“ear”
oosten	1,676	6	*z − r − x*	[*mizrax*]	5,362	0	8	“east”
opdracht	3,722	13	*q − d − ṡ*	[*haqdaṡa*]	215	11	3	“dedication”
opgraving	110	1	*x − p − r*	[*xapira*]	260	1	11	“excavation”
orthodox	382	3	*x − r−d*	[*xaredi*]	1,224	7	4	“orthodox”
ouderdom	538	11	*z − q − n*	[*ziqna*]	1,299	3	8	“old age”
overrijden	95	2	*d − r − s*	[*daras*]	214	1	7	“to run over(a car)”
parabel	42	1	*m − ṡ − l*	[*maṡal*]	3,759	5	7	“parable”
paramedisch	43	1	*x − b − ḣ*	[*xobeš*]	747	1	7	“medical assistant”
ploegende	0	25	*x − r − ṡ*	[*xariš*]	445	10	7	“ploughing”
poort	1,468	28	*ṡ − ‘ − r*	[*ṡa‘ar*]	69,710	5	2	“gate”
praten	14,570	53	*d−b − r*	[*diber*]	6,349	5	14	“to talk”
preek	379	14	*d − r − ṡ*	[*draṡa*]	1,407	5	9	“sermon”
print	3	3	*d − p − s*	[*dpus*]	1,877	0	11	“print”
puber	170	5	*b−g−r*	[*mitbager*]	523	0	8	“teenager”
racisme	110	0	*g − z −‘*	[*giz‘anut*]	1,253	0	6	“racism”
regelen	2,497	123	*s − d − r*	[*sider*]	993	0	22	“to arrange”
regen	2,336	58	*g − ṡ − m*	[*gešem*]	5,190	8	2	“rain”
regering	4,887	90	*m − ṡ − l*	[*memšala*]	25,800	7	5	“government”
reglement	329	9	*t − q − n*	[*taqanon*]	811	11	7	“reglament”
reptiel	136	2	*z − x − l*	[*zoxel*]	175	0	11	“reptile”
rijpheid	120	3	*b − ṡ − l*	[*bšelut*]	586	11	3	“maturation”
schandalig	113	1	*x − p − r*	[*maxpir*]	382	10	2	“shameful”
scheiding	1,589	27	*p − l − g*	[*pilug*]	161	3	13	“split”
schrijven	21,201	135	*x − b − r*	[*xiber*]	783	10	9	“to write”
schroef	283	18	*b − r − g*	[*boreg*]	7,990	0	9	“screw”
sluiting	314	27	*s − g − r*	[*sgira*]	1,594	7	15	“closing”
snelheid	1,694	26	*m − h − r*	[*mhirut*]	4,555	0	7	“speed”
spanning	3,164	38	*m − t − x*	[*metax*]	5,873	0	13	“tension”
springen	5,030	49	*z − n − q*	[*zineq*]	1,196	0	5	“to jump”

(*continued on next page*)

Appendix A (*continued*)

Dutch			Hebrew					English
Word	Freq.	Fam. size	Root	Word	Freq.	Unrelated	Related	Word
staande	17	0	*d – g – l*	[*dagal*]	20,252	5	12	“to stand for”
steunzool	14	1	*d – r – s*	[*midras*]	69	7	1	“sole”
stof	4,319	139	*ḣ – b – q*	[*ḣabaq*]	2,266	4	12	“dust”
struikelen	790	5	*m – ‘– d*	[*ma‘ad*]	108	0	3	“to stumble”
tak	2,305	23	*s – b – k*	[*sbak*]	364	7	3	“thick branch”
telegram	399	4	*b – r – q*	[*mibraq*]	2,509	5	2	“telegram”
tent	1,141	37	*ḣ – h – l*	[*ḣohel*]	1,580	2	2	“tent”
tik	387	13	*T – p – x*	[*Tpixa*]	86	4	2	“tap”
toegewijd	138	1	*d – b – q*	[*dabeq*]	2,217	14	3	“devoted”
trede	344	2	*d – r – g*	[*madrega*]	2,057	0	12	“step”
treurig	853	1	*š – ‘ – r*	[*ši‘er*]	13	1	5	“sad”
tumor	281	7	*g – d – l*	[*gidul*]	10,717	16	5	“tumor”
vaart	1,025	89	*p – l – g*	[*haplaga*]	494	13	3	“sailing”
vangen	3,435	31	*l – k – d*	[*lakad*]	129	8	9	“to capture”
verenigen	3,905	59	*l – k – d*	[*hitlaked*]	23	9	8	“become united”
vergeven	1,028	14	*m – x – l*	[*maxal*]	49	0	4	“to forgive”
verhaal	10,107	23	*s – p – r*	[*sipur*]	28,529	20	16	“story”
verkoper	614	3	*z – b – n*	[*zaban*]	68	0	2	“salesman”
verkorting	99	3	*q – š – r*	[*qišur*]	1,833	5	15	“shortening”
verlies	1,765	15	*ḣ – b – d*	[*ṡabeda*]	433	3	8	“loss”
verraden	1,406	8	*b – g – d*	[*bgida*]	1,044	2	7	“treason”
vertraging	367	3	*ḣ – x – r*	[*ḣixur*]	381	3	6	“delay”
vervullen	3,189	10	*g – ṡ – m*	[*higṡim*]	236	2	8	“to perform”
verzegelen	161	2	*ḣ – T – m*	[*ḣaTam*]	23	0	11	“to seal”
vlag	1,217	13	*d – g – l*	[*degel*]	2,553	2	3	“flag”
vluchten	1,836	51	*b – r – x*	[*barax*]	2,828	1	8	“to run away”
voedsel	2,792	46	*ḣ – k – l*	[*ḣoxel*]	33,359	0	11	“food”
vriend	12,053	59	*x – b – r*	[*xaber*]	24,814	9	10	“friend”
vuilnisbelt	70	1	*z – b – l*	[*mizbala*]	26	4	3	“garbage site”
wagen	5,148	104	*‘– g – l*	[*‘agala*]	538	12	3	“wagon”
wapen	3,002	60	*n – ṡ – q*	[*neṡeq*]	7,489	12	4	“weapon”
wassen	2,475	93	*r – x – š*	[*hitraxeš*]	10	0	9	“wash oneself”
wassend	0	93	*k – b – s*	[*kbisa*]	2,193	0	9	“laundering”
wijnstok	180	1	*g – p – n*	[*gepen*]	5,455	0	1	“vine”
woede	2,755	14	*z – ‘– m*	[*za‘am*]	3,966	1	6	“rage”
zaad	1,339	40	*z – r –‘*	[*zera‘*]	1,432	0	7	“seed”
zeggen	148,830	45	*ḣ – m – r*	[*ḣamar*]	140,962	0	8	“to say”
zelfmoord	679	9	*ḣ – b – d*	[*hitḣabed*]	745	8	3	“committed suicide”
zon	1,953	93	*ṡ – m – ṡ*	[*ṡemeṡ*]	6,719	12	3	“sun”

For simplicity we have not transcribed the geminations or “doublings” of root consonants in the Hebrew transcriptions.

References

Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language, 40*, 41–61.

Aronoff, M. (1994). *Morphology by itself: Stems and inflectional classes*. Cambridge, Mass.: The MIT Press.

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

Baayen, R. H. (2005). Data mining at the intersection of psychology and linguistics. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones*. Hillsdale, NJ: Erlbaum.

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

Baayen, R. H., Moscoso del Prado, F., Schreuder, R., & Wurm, L. (2003). When word frequencies do NOT regress towards the mean. In R. H. Baayen & R. Schreuder (Eds.), *Morphological structure in language processing* (pp. 463–484). Berlin: Mouton de Gruyter.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (CD-ROM), linguistic data consortium*. Philadelphia, PA: University of Pennsylvania.

Baayen, R. H., Tweedie, F. J., & Schreuder, R. (2002). The subjects as a simple random effect fallacy: Subject variability and morphological family effects in the mental lexicon. *Brain and Language, 81*, 55–65.

Bates, E., D'Amico, S., Jacobsen, T., Szekely, A., Andonova, E., Devescovi, A., et al. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review, 10*(2), 344–380.

Becker, C. A. (1979). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 252–259.

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics. Identifying influential data and sources of collinearity*. Wiley, New York: Wiley Series in Probability and Mathematical Statistics.

Berman, R. (1978). *Modern Hebrew structure*. Tel Aviv: University Publishing Projects.

Bertram, R., Schreuder, R., & Baayen, R. H. (2000). The balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 419–511.

Bolle, M., & Pimentel, J. (1984). *Woordenboek Nederlands Hebreeuws*. Naarden, The Netherlands: Strengholt's Boeken.

Borowsky, R., & Besner, D. (1993). Visual word recognition: A multistage activation model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 813–840.

Bradley, D. C., & Forster, K. I. (1987). A reader's view of listening. *Cognition, 25*, 103–134.

Carroll, J. B. (1967). On sampling from a lognormal model of word frequency distribution. In H. Kučera & W. N. Francis (Eds.), *Computational analysis of present-day American english* (pp. 406–424). Providence: Brown University Press.

Chatterjee, S., Hadi, A., & Price, B. (2000). *Regression analysis by example*. New York: John Wiley.

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

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

De Jong, N. H., Feldman, L. B., Schreuder, R., Pastizzo, M., & Baayen, R. H. (2002). The processing and representation of Dutch and English compounds: Peripheral morphological, and central orthographic effects. *Brain and Language, 81*, 555–567.

De Jong, N. H., Schreuder, R., & Baayen, R. H. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes, 15*, 329–365.

Deutsch, A., Frost, R., Pollatsek, A., & Rayner, K. (2000). Early morphological effects in word recognition in Hebrew: Evidence from parafoveal preview benefit. *Cross-linguistic perspectives on morphological processing, 15*(4–5), 487–506.

Dijkstra, T., Moscoso del Prado Martín, F., Schulpen, B., Schreuder, R., & Baayen, R. (2005). A roommate in cream: Morphological family size effects on interlingual homograph recognition. *Language and Cognitive Processes, 20*(1), 7–42.

Downie, N. M., & Heath, R. W. (Eds.). (1965). *Basic statistical methods*. New York: Harper and Row.

Feldman, L. B., & Pastizzo, M. J. (2003). Morphological facilitation: the role of semantic transparency and family size. In R. H. Baayen & R. Schreuder (Eds.), *Morphological structure in language processing* (pp. 233–258). Berlin: Mouton de Gruyter.

Feldman, L. B., & Siok, W. W. T. (1997). The role of component function in visual recognition of Chinese characters. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23*, 778–781.

Feldman, L. B., & Siok, W. W. T. (1999). Semantic radicals contribute to the visual identification of Chinese characters. *Journal of Memory and Language, 40*, 559–576.

Feldman, L. B., Frost, R., & Pnini, T. (1995). Decomposing words into their constituent morphemes: Evidence from English and Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 947–960.

Frost, R., Deutsch, A., & Forster, K. (2000). Decomposing morphologically complex words in a nonlinear morphology. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(3), 751–765.

Frost, R., Deutsch, A., Gilboa, O., Tannenbaum, M., & Marslen-Wilson, W. D. (2000). Morphological priming: Dissociation of phonological, semantic and morphological factors. *Memory & Cognition, 28*(8), 1277–1288.

Frost, R., Forster, K. I., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 829–856.

Gaskell, M. G., & Marslen-Wilson, W. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes, 12*, 613–656.

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

Landauer, T., & Dumais, S. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological Review, 104*(2), 211–240.

Lorch, R. F., & Myers, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 149–157.

Lüdeling, A., & De Jong, N. H. (2002). German particle verbs and word-formation. In N. Dehé, R. Jackendoff, A. McIntyre, & S. Urban (Eds.), *Verb–particle explorations* (pp. 315–333). Berlin: Mouton de Gruyter.

McCarthy, J. J. (1981). A prosodic theory of non-concatenative morphology. *Linguistic Inquiry, 12*, 373–418.

Morton, J. (1969). The interaction of information in word recognition. *Psychological Review, 76*, 165–178.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS, statistics and computing*. New York: Springer.

Plaut, D. C., & Booth, J. R. (2000). Individual and developmental differences in semantic priming: Empirical and computational support for a single mechanism account of lexical processing. *Psychological Review, 107*, 786–823.

Pylkkänen, L., Feintuch, S., Hopkins, E., & Marantz, A. (2004). Neural correlates of the effects of morphological family frequency and family size: An MEG study. *Cognition, 91*, B35–B45.

Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language, 37*, 118–139.

Schweika, Y. (1997). *Modern Hebrew dictionary*. Tel-Aviv: Rav-Milim Matach.

Stanovich, K. E., & West, R. F. (1981). The effect of sentence context on ongoing word recognition: Test of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 658–672.

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

Taft, M., & Zhu, X. (1997). Submorphemic processing in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 761–775.