Morphological Family Size in a Morphologically Rich Language: The Case of Finnish Compared With Dutch and Hebrew

Fermín Moscoso del Prado Martín
Medical Research Council Cognition and Brain Sciences Unit, University of Nijmegen, and Max Planck Institute for Psycholinguistics

Raymond Bertram and Tuomo Häikö
University of Turku

Robert Schreuder and R. Harald Baayen
University of Nijmegen and Max Planck Institute for Psycholinguistics

Finnish has a very productive morphology in which a stem can give rise to several thousand words. This study presents a visual lexical decision experiment addressing the processing consequences of the huge productivity of Finnish morphology. The authors observed that in Finnish words with larger morphological families elicited shorter response latencies. However, in contrast to Dutch and Hebrew, it is not the complete morphological family of a complex Finnish word that codetermines response latencies but only the subset of words directly derived from the complex word itself. Comparisons with parallel experiments using translation equivalents in Dutch and Hebrew showed substantial cross-language predictivity of family size between Finnish and Dutch but not between Finnish and Hebrew, reflecting the different ways in which the Hebrew and Finnish morphological systems contribute to the semantic organization of concepts in the mental lexicon.

In languages such as English and Dutch, stems differ in their productivity. Some stems give rise to a great many complex words. For instance, in English, the stem man appears in nearly 200 complex words. Other stems hardly ever give rise to complex words, for example, the noun scythe, which only has its corresponding verb (to scythe) as a morphological relative. Previous research has shown that the morphological family size of a stem, defined as the number of different complex words in which the stem appears as a constituent, is a robust predictor of response latencies in tasks such as visual lexical decision, auditory lexical decision, and subjective familiarity rating. Words with a larger morphological family size elicit shorter response latencies and higher subjective familiarity scores than do words with smaller family sizes matched for frequency (Schreuder & Baayen, 1997).

The effect of family size is present when measures of word form such as orthographic neighborhood size and bigram frequency are controlled for (Schreuder & Baayen, 1997). Family-size counts are highly correlated with measures of morpheme frequency (Baayen, Tweedie, & Schreuder, 2002; Schreuder & Baayen, 1997). Both factorial studies (De Jong, Schreuder, & Baayen, 2000; Schreuder & Baayen, 1997) and regression studies (Baayen et al., 2002) have shown that family-size effects can be observed independent from the effects of morpheme frequency. The family-size effect is present independently of morpho–phonological inconsistency (De Jong et al., 2000) and remains a significant predictor when age of acquisition is partialed out (De Jong, 2002; see also Moscoso del Prado Martín, 2003).

The effect of morphological family size is well-established for Germanic languages (Dutch: Bertram, Baayen, & Schreuder, 2000; De Jong et al., 2000; Schreuder & Baayen, 1997; English: De Jong, Feldman, Schreuder, Pastizzo, & Baayen, 2002; German: Lüdeling & De Jong, 2002). Recently, an effect of morphological family size has also been established for a Semitic language, Hebrew (Moscoso del Prado Martín et al., 2003). In this language, morphological family size is defined in terms of the number of words that share a given consonantal root. The morphological family size in Hebrew ranges between 1 and 30 and is therefore much more restricted than English family sizes (range = 1–200) and Dutch family sizes (range = 1–550). Even though morphological families tend to be small in Hebrew, morphological family...
size emerged as a reliable predictor of response latencies independently of word frequency.

The family-size effect is semantic in nature (Bertram et al., 2000; De Jong, 2002). Recent evidence supporting this conclusion has been obtained in Hebrew as well as for Dutch–English bilinguals. The Hebrew family-size effect has a specific property that is particular to the Hebrew root, namely, that for words with homonymic roots, the semantically related family members lead to facilitation whereas the semantically unrelated family members give rise to inhibition. In Dutch, such an effect has not been observed for homonymic stems (De Jong, 2002). However, a similar effect has been observed for interlingual homographs for Dutch–English bilinguals (Dijkstra, Moscoso del Prado Martín, Schulpen, Schreuder, & Baayen, in press). Interlingual homographs are noncognate words with identical spelling but different meanings across two languages. For instance, angel refers to a celestial being in English and to a sting of a bee or wasp in Dutch. When Dutch bilinguals performed a Dutch simple visual lexical decision task with interlingual homographs as target words, the number of Dutch family members of the interlingual homographs is negatively correlated with response latencies (facilitation) whereas the number of English family members is positively correlated with the response latencies (inhibition). In contrast, when participants performed English simple lexical decision, the same homographs elicited response latencies that correlated positively with the Dutch family counts and negatively with the English counts. Because interlingual homographs such as angel have different meanings in the two languages, the opposite effects of family size observed for the Dutch and English family sizes support the hypothesis that the morphological family size arises at semantic levels of lexical processing.

It is important to realize that the family-size measure explores semantic relations between sets of words (see, e.g., Moscoso del Prado Martín, Kostić, & Baayen, 2004), whereas virtually all other studies known to us focus on semantic relations between pairs of words, such as synonymy, hyponymy, and hyperonymy, or associative measures obtained from ratings (e.g., McRae, DeSa, & Seidenberg, 1997). In fact, the family-size effect turns out to be a much stronger predictor than such measures; see, for example, the reanalysis of the data of McRae et al. (1997) provided by De Jong, Schreuder, and Baayen (2003). This greater predictivity is not so much stronger than that of such measures; see, for example, the reanalysis of the data of McRae et al. (1997) provided by De Jong, Schreuder, and Baayen (2003). This greater predictivity is not so surprising once it is realized that a great many semantic relations in the lexicon are expressed morphologically. Furthermore, semantic categorization judgments and latencies are co-determined by morphological family size (De Jong, 2002). Finally, an information-theoretic account of the family-size effect (in which the token frequencies of the family members are also considered) can be found in Moscoso del Prado Martín et al. (2004). This study shows that the family-size count is a very good estimate of the informational complexity of morphological paradigms.

In this study, we report an experiment addressing the possible existence of a family size in Finnish. Finnish belongs to the Finno–Ugric language family and is well known for its rich and complex morphology. It combines a complex inflectional system with a great many cases with productive derivation involving rampant stem allomorphy and very productive compounding. In Finnish, a stem such as työ [work] has roughly 7,000 family members, including työntekijä [employee], työelämisopimus [wage rate treaty], työstökö [machine tool], työläs [laborious], and työväenluokka [working class]. Obviously, most Finnish stems have smaller morphological families, but many are very sizeable anyway, with family sizes of some 200 words or more.

Whereas the Hebrew study established that family-size effects generalize from Germanic concatenative morphology to Semitic nonconcatenative morphology, in the present study we investigate whether family-size effects also exist in a language with agglutinative1 morphology like Finnish. It is far from evident that this would be the case. Like the word-frequency effect, the family-size effect is logarithmic in nature. Robust effects are typically observed in the range of 0–40 family members, after which there is generally a floor effect. Given the large families counted for Finnish stems, no effect of family size might be observed because of an overall floor effect. As we show below, this prediction is partially correct, requiring a more limited family-size definition for complex words.

In a previous study, Moscoso del Prado Martín et al. (2003) performed two lexical decision experiments in Hebrew and Dutch. The materials for their Dutch experiment were the Dutch translations of the words that were used in the Hebrew experiment. Their results showed that Hebrew response latencies can be predicted from the Dutch family sizes of the corresponding translation equivalents even after Hebrew frequency and Hebrew family size are partialled out and vice versa. This indicates that there is substantial similarity in semantic lexical organization in Dutch and Hebrew, even though these languages are typologically fundamentally different.

A second question addressed in the current study is whether a similar cross-language predictivity might be observed for Finnish and Dutch translation equivalents and for Finnish and Hebrew translation equivalents. The patterns of cross-language predictivity have important implications for the degree of isomorphy in semantic organization across languages with typologically different morphological systems.

The purpose of the present study can therefore be summarized as follows. First, by attesting the role of a family-size effect in Finnish, we provide further validation for this measure. Second, by means of cross-linguistic comparisons, we further illustrate the potential of this measure as a research tool for investigating semantic organization in the mental lexicon.

The following visual lexical decision experiment addresses the questions raised above. It is designed along the lines of the Hebrew and Dutch experiments reported by Moscoso del Prado Martín et al. (2003) and makes use of translation equivalents of the Hebrew and Dutch words used in that study.

**Experiment**

**Method**

*Participants.* Twenty-six undergraduate students of the University of Turku participated in the experiment. All were native speakers of Finnish and had normal or corrected-to-normal vision.

*Materials.* The materials of these studies are the translation equivalents of the Hebrew and Dutch words used in the experiments reported in

1 Agglutinative languages are languages with an especially rich concatenative morphology in which grammatical functions tend to be expressed by separate affixes. For instance, the Finnish word Taloissaninkinokko translates into the English sentence “(Do you mean) in my houses, too?” and consists of the morphemes talo-i:ssa-ni-kin-ko.
Moscoso del Prado Martín et al. (2003). As our point of departure, we took the 162 Dutch words from their Experiment 2 and translated them into Finnish. The translations were done using a Dutch–Finnish dictionary (Schouwvlieger & Mäkinen-Schouwvlieger, 1992), and they were extensively validated by Raymond Bertram and Tuomo Häikiö. When a word had different possible translations into Finnish with different meanings, we included all translation possibilities in the experiment. Four of the original Dutch words could be translated into Finnish only by using multiword utterances and were excluded from the experiment. In this way, we obtained a set of 167 Finnish words. Of these words, 71 were morphologically simple, 81 were derived words, and 15 were compounds. Within the derived words, there are only 3 that contain a suffix for which a base frequency effect has ever been reported in the literature (Bertram, Baayen, Schreuder, Laine, & Hyöna, 2000; Järvikivi, Bertram, & Niemi, 2003; Vannest, Bertram, Järvikivi, & Niemi, 2002).

Frequency counts for these words are based on the unpublished computerized Turun Sanomat Finnish newspaper corpus of 22.7 million word forms accessed with the help of the WordMill database program of Laine and Virtanen (1999). Morphological family-size counts were also based on this database, with each of the potential family members evaluated by Tuomo Häikiö, in some occasions aided by a dictionary (Nykyajan sanakirja, 1978). Each of these words was paired with a pseudoword whose phonotactics did not violate the phonology of Finnish. The pseudowords were derived from the experimental target words by changing 2–7 characters. Monomorphemic words were predominantly changed into pseudowords without any morphological structure (e.g., 2–7 characters. Monomorphemic words were predominantly changed into pseudowords with a real stem but no suffix (e.g., Tuomo Häikiö, in some occasions aided by a dictionary (Nykyajan sanakirja, 1978). Each of these words was paired with a pseudoword whose phonotactics did not violate the phonology of Finnish. The pseudowords were derived from the experimental target words by changing 2–7 characters. Monomorphemic words were predominantly changed into pseudowords without any morphological structure (e.g., jalka [leg] was transformed into solka, varas [thief] became taras); occasionally, we created pseudowords with a real stem but no suffix (e.g., vaatua [to take care] became puula, which contains the stem pau [tree]). For derived words, only the stem was altered into a pseudostem, but the suffix remained intact (e.g., hävytön [shameless] became selytön, in which -tön corresponds to “-less”). For compound words, sometimes the first or second constituent was replaced by an alternative existing constituent (e.g., first constituent: lampunvarjostin [lampshade] was transformed into lennonvormostin, in which lennon means “flight” in general; second constituent: nenärenkas [nose ring] became nypäkangas, in which kangas means “textile”); sometimes the whole word was changed so that no sublexical morphological structure was present anymore (e.g., itsemurha [self murder, suicide] became istekorha).

Twenty practice trials, 10 words and 10 pseudowords, were run before the actual experiment. We constructed three different permutations and their corresponding reversed versions of the original word list for counterbalancing. Table 1 provides a summary of the distributional properties of the data set.

Procedure. Participants were tested 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 Finnish word. Following a pause after the test trials, the experiment was run with two further pauses, dividing the experiment into three blocks, each containing one third of the materials. Items were preceded by a fixation mark in the middle of the screen for 500 ms. After 500 ms, the stimulus appeared at the same position. Stimuli were presented in white lowercase 12-point Helvetica letters on a dark background, and they remained on the screen for 1,500 ms. The maximum time span allowed for a response was 2,000 ms from stimulus onset.

Results and Discussion

All participants in this experiment performed with an error rate of less than 15%. One item elicited errors for more than 30% of the participants and was thus excluded from the analyses. Additionally, we excluded four items that elicited response latencies of more than two and a half standard deviations above or below the mean.

Table 1 provides the medians, means, standard deviations, and ranges for the frequency, family-size, and word-length counts for this data set and the average response latencies in the experiment after excluding the four outliers. In addition, it lists details about the error scores. As the analyses of the error data revealed the same pattern of results as the reaction times, separate analyses of the errors are not reported.

A multilevel regression model (Baayen et al., 2002; Pinheiro & Bates, 2000; a more powerful extension of the technique described by Lorch & Myers, 1990) fit to the dataset, with log response latency as the dependent variable and log frequency, log family size, and word length as the independent variables, revealed a facilitatory main effect for word frequency, $F(1, 3625) = 521.86$, $p < .0001$; an inhibitory main effect of word length, $F(1, 3625) = 137.66$, $p < .0001$, after the effect of frequency was partialed out; and a facilitatory main effect of family size, $F(1, 3625) = 24.62$, $p < .0001$, after the effects of frequency and word length were partialed out. We also observed a significant interaction between word length and word frequency, $F(1, 3625) = 89.21$, $p < .0001$, after partiauling out the main effects: Longer words elicited longer response latencies but only for lower frequency words.

These results document, for the first time, the presence of a morphological family-size effect in Finnish. As in English, German, and Dutch as well as in Hebrew, words with larger families give rise to shorter response latencies than do words with smaller families. The presence of a morphological family-size effect in three genetically unrelated language families, Indo–European, Hamo–Semitic, and Finno–Ugric, shows that, across typologically very different morphological systems, the organization of related words in morphological paradigms (i.e., the set of all the words that share a given morphological constituent) is an important factor in lexical processing.

Thus far, it would seem that the possibility we considered in the introduction, namely, that the large family sizes of Finnish compared with English or Dutch would lead to a floor effect, is not borne out. However, consider the selection of the members of the morphological family of kirja [book] in Finnish shown in the Appendix. Note that although there is a family member that has a translation in English that contains the stem book (notebook), all other family members require translations with quite different stems in English, ranging from author to library and from register to dissertation. Note furthermore that some family members form semantically cohesive clusters, such as the words for library, librarian, and public library. This suggests the possibility that the family-size effect in Finnish might be carried predominantly or perhaps even exclusively by the semantically more closely related family members.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mdn</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>670</td>
<td>3,155</td>
<td>7,097</td>
<td>1–56,193</td>
</tr>
<tr>
<td>Word length</td>
<td>7</td>
<td>7.2</td>
<td>2.3</td>
<td>3–14</td>
</tr>
<tr>
<td>Family size</td>
<td>298</td>
<td>620</td>
<td>892</td>
<td>8–6,029</td>
</tr>
<tr>
<td>Dominated family size</td>
<td>88</td>
<td>273</td>
<td>485</td>
<td>0–3,080</td>
</tr>
<tr>
<td>Nondominated family size</td>
<td>29</td>
<td>347</td>
<td>762</td>
<td>0–5,835</td>
</tr>
<tr>
<td>Response latency (ms)</td>
<td>604</td>
<td>617</td>
<td>63</td>
<td>530–808</td>
</tr>
<tr>
<td>Error rate</td>
<td>.000</td>
<td>.017</td>
<td>.033</td>
<td>.000–.192</td>
</tr>
</tbody>
</table>
One way of obtaining an objective and replicable way of defining the notion of being more closely related semantically is to make a distinction between the family members of a word that are its direct descendants (its dominated family) and the other family members (its nondominated family). Figure 1 illustrates the distinction between the dominated and nondominated family size for the Finnish family of työläinen [worker]. The dominated family size of työläinen consists of the words that are shown in bold in the figure. Its nondominated family size consists of the remaining words. Note that the dominated family members are in general more closely related in meaning to each other than is the case for the nondominated family members. This leads to the hypothesis that in Finnish, the morphological family size might be carried predominantly or perhaps exclusively by the dominated family size.

To test this hypothesis, we selected the 83 complex Finnish words in our data set. (We excluded the monomorphemic words from the analysis, as for monomorphemic words the family size as a whole is identical to the dominated family size, with the nondominated family size being an empty set.) For these complex words, we determined the dominated and nondominated family size. We then carried out a regression analysis, with log word frequency, word length, log dominated family size, and log nondominated family size as the independent variables and log response latencies as the dependent variable. A multilevel regression analysis revealed a highly significant effect for the dominated family size, $F(1, 2127) = 20.25, p < .0001$, and no effect whatsoever for the nondominated family size ($F < 1$). In fact, it turns out that the total family size is not a good predictor for the complex words in our data. This shows that adding the nondominated family members to the family-size count for complex words in Finnish amounts to adding so much noise that the effect of the true predictor, the dominated family size, is completely masked.

The nonexistence of a family-size effect for the nondominated family is partly in line with the intuition outlined in the introduction that with the large family sizes of Finnish, the family-size effect might be reduced because of a floor effect. However, restriction of the effect to the dominated family suggests that the degree of semantic relatedness in the family might be the key determinant rather than size as such. To gain further insight into the weight of these two factors, the magnitude of the family on the one hand and its semantic cohesion on the other, we reanalyzed the Dutch analogue of the present experiment reported in Moscoso del Prado Martín et al. (2003), in which the translation equivalents of the Finnish words studied in the present article were analyzed. From their Dutch items, we selected the 59 words that were morphologically complex. A multilevel regression model revealed significant effects of both dominated and nondominated family size, although the beta weight for the dominated family size, $\hat{\beta} = -0.085, SE = .017, t(2018) = -4.875, p < .0001$, was more than twice as large as the beta weight of the nondominated family size, $\hat{\beta} = -0.030, SE = .011, t(2018) = -2.790, p = .0053$.

This result suggests that the dominated family size is the prime carrier of the family size but that the nondominated family size may also have some predictive power, at least in Dutch. This is probably due to the relatively small sizes (at least compared with Finnish) of morphological families in Dutch. Within these small families, there is enough semantic similarity between the nondominated and the dominated family members to allow a nondominated family-size effect to emerge. In Finnish, by contrast, the range of meanings covered by the nondominated family is too broad, leading to semantic neighborhoods that are too sparsely populated to give rise to a measurable family-size effect in the response latencies.

At this point, it should be made explicit that we do not claim that the distinction between the dominated and the nondominated family is an absolute distinction for Finnish. To the contrary, we believe that closely related nondominated family members will also contribute to the family-size effect. However, we leave it to
further research to establish principled ways in which the contrib-
uting nondominated family members might be ascertained.

In summary, the crucial contribution of the present experiment
to our knowledge of the family-size effect in human cognition is
that, through the examination of the family-size effect in a highly
productive agglutinative language such as Finnish, the semantic
nature of the effect is clarified in more detail. If the family-size
effect were just a form effect, the distinction between the domi-
nated and nondominated family should not have been relevant,
contrary to fact. This shows that the family-size effect depends on
the combination of shared morphological form and shared seman-
tics. When the condition of semantic overlap is not met, as for
most nondominated family members in Finnish, those family
members no longer contribute to the effect.

Cross-Language Analyses

As mentioned in the introduction, Moscoso del Prado Martín et
al. (2003) observed that Hebrew response latencies can be pre-
pdicted from the Dutch family sizes of the corresponding translation
equivalents even after Hebrew frequency and Hebrew family size
have been partialed out first and vice versa. This result is indicative
of substantial similarity in semantic lexical organization in Dutch
and Hebrew, even though these languages are fundamentally dif-
ferent typologically. We now investigate whether a similar cross-
language predictivity might be observed for Finnish and Dutch
translation equivalents and also for Finnish and Hebrew translation
equivalents. This allows us to obtain insight in the extent of
cross-language predictivity across typologically unrelated lan-
guages and its implications for the degree of isomorphism in seman-
tic organization across radically different morphological systems.

For this cross-language multiple regression analysis, we selected
those items that elicited less than 30% errors in the three experi-
ments in Hebrew, Dutch, and Finnish. In this way, we obtained a
total of 131 items, each with three response latencies. For each
word in each of the three languages, we added as predictors length
(in letters), word frequency, and morphological family size in that
language. The key question of interest is whether length, fre-
quency, and family size of, for example, Dutch would predict
response latencies in Finnish, even after the effects of Finnish
frequency, Finnish word length, and Finnish family size have been
partialed out.

Table 2 summarizes the results obtained for the six pairwise
comparisons (Hebrew with Dutch, Hebrew with Finnish, Dutch
with Hebrew, Dutch with Finnish, Finnish with Hebrew, and
Finnish with Dutch). When predicting from Language A to Lan-
guage B, we took the best multilevel regression model fitted to the
data from Language B as point of departure. The columns of Table
2 list the language for which the response latencies are predicted.
The rows of Table 2 list the language from which an additional
predictor (frequency, length, or family size) is taken. Each F
statistic and associated p value correspond to a separate analysis
including the within-language variables and one additional predic-
tor from another language. (Including more than one additional
predictor at a time would have led to a serious collinearity prob-
lem.) For the details of the within-language regression models for
Hebrew and Dutch, the reader is referred to Moscoso del Prado
Martín et al. (2003). For the Finnish data, the within-language
model incorporates the effects of word frequency, length in letters,
and family size: the full family size for the monomorphemic words
and the dominated family size for the complex words. The third
row of Table 2 shows that Finnish frequency is an excellent
predictor of Dutch response latencies, after the effects of Dutch
frequency, length, and family size have been partialed out. Finnish
family size likewise emerged as a highly significant predictor, and
even Finnish length turned out to have some predictive value.

What Table 2 shows is that frequency is an excellent additional
predictor in five out of six cases. The only instance in which
frequency fails to have additional predictivity is when Finnish
frequency is used to predict Hebrew reaction times. Note that, in
terms of stem productivity, the typological distance is greatest
between Hebrew and Finnish, with Dutch taking an intermediate
position. Family size emerges alongside word frequency as a

Table 2
Cross-Language Predictivity of Word Frequency, Word Length, and Morphological Family Size
Between Translation Equivalents in Hebrew, Dutch, and Finnish, in Sequential Analyses of
Variance in Multilevel Regression Analyses

<table>
<thead>
<tr>
<th>Language and predictors</th>
<th>Hebrew</th>
<th>Dutch</th>
<th>Finnish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hebrew Frequency</td>
<td></td>
<td>F(1, 4603) = 7.39***</td>
<td>F(1, 3263) = 20.57***</td>
</tr>
<tr>
<td>Word length</td>
<td></td>
<td>F(1, 4603) = 2.53</td>
<td>F(1, 3263) = 11.51***</td>
</tr>
<tr>
<td>Related family size</td>
<td></td>
<td>F(1, 4603) = 20.22***</td>
<td>F(1, 3263) = 1.38</td>
</tr>
<tr>
<td>Dutch Frequency</td>
<td>F(1, 3184) = 15.28***</td>
<td></td>
<td>F(1, 3263) = 35.99***</td>
</tr>
<tr>
<td>Word length</td>
<td>F(1, 3184) = 6.17*</td>
<td></td>
<td>F(1, 3263) = 56.94***</td>
</tr>
<tr>
<td>Family size</td>
<td>F(1, 3184) = 15.03***</td>
<td></td>
<td>F(1, 3263) = 18.62***</td>
</tr>
<tr>
<td>Finnish Frequency</td>
<td>F(1, 3184) = 2.20</td>
<td>F(1, 4603) = 15.98***</td>
<td></td>
</tr>
<tr>
<td>Word length</td>
<td>F &lt; 1</td>
<td>F(1, 4603) = 4.90*</td>
<td></td>
</tr>
<tr>
<td>Family size</td>
<td>F(1, 3184) = 1.52</td>
<td>F(1, 4603) = 16.75***</td>
<td></td>
</tr>
</tbody>
</table>

Note. The columns specify the language for which the response latencies are predicted, and the rows indicate
the languages from which the independent variables are taken. The value listed in a given cell specifies the
significance of the predictor listed in the row of the cell after the within-language effects of frequency, length,
and family size are partialed out.

* p < .05. ** p < .005. *** p < .0005.
remarkable explanatory variable in four out of six cases. The two cases in which family size fails as a cross-language predictor is from Finnish family size to Hebrew reaction times and from Hebrew family size to Finnish reaction times. Again, cross-language predictivity breaks down where the typological difference in morphological structure and stem productivity is greatest. Finally, even word length shows some cross-language predictivity. The only language pair for which word length is predictive in both directions is Finnish and Dutch. The small differences in word length in Hebrew seem not to be predictive for Dutch but predictive for Finnish. Conversely, the big differences in word lengths in Finnish emerge as predictive for Dutch but not for Hebrew.

Figure 2 summarizes the family-size effects for the three languages by means of nonparametric regression lines. Note that the effect of family size is greatest for lower family sizes across all three languages and that it levels off for greater family counts. The point of inflection is different for the three languages. For Hebrew, it is around 7.4 ($e^2$); for Dutch, it is around 20.1 ($e^3$); and for Finnish, it is around 148 ($e^5$). Although, in comparison to Dutch, Finnish morphological families are very large and thus more sensitive to floor effects. Therefore, it is possible that we could be observing a floor effect in the Finnish analyses.

General Discussion

The questions addressed in this study were, first, whether the family-size effect might be observed in Finnish and, second, to what extent Finnish might participate in the cross-language predictivity of family size observed for Hebrew and Dutch. As to the first question, a visual lexical decision experiment revealed that, as in Germanic languages such as Dutch, English, and German as well as in Hebrew (Semitic), the morphological family size is also relevant for lexical processing in Finnish, a Finno–Ugric language. This finding provides further evidence for the cross-linguistic generality of the family-size effect.

Earlier studies (De Jong, 2002; Moscoso del Prado Martín et al., 2003) established that the observed effect of the morphological family size probably arises at the level of semantic processing. These studies also established that semantic similarity shared among the family members is crucial for the effect to emerge.\(^2\) Inspection of morphological families in Finnish, however, suggests that the larger families as a whole are semantically fairly diverse. To obtain further insight into the role of semantic similarity, we introduced the notion of the dominated versus the nondominated family size for complex words. The dominated family size (consisting of the semantically more similar morphological descendants of a complex word) turned out to be the crucial

\(^2\) In this study, we have used word length as a means for assessing the effect of family size while controlling for an important variable relating to word form. Other measures, such as neighborhood size and orthographic bigram frequency, were not included in our analyses for two reasons. First, Schreuder and Baayen (1997) showed that the family-size effect is confounded neither with neighborhood size nor with orthographic bigram frequency. Second, adding such measures to the regression models leads to a very large increase in collinearity, with the condition number (Belsley, 1991; Belsley, Kuh, & Welsch, 1980) increasing from 12 to 79 when just bigram frequency is added. With such high collinearity the coefficients of the regression model become unstable. When a regression model with bigram frequency as an additional predictor is nevertheless fitted to the response latencies of the Dutch experiment, the effects of frequency, word length, and family size remain highly significant.
predictor for Finnish. A reanalysis of Dutch data showed both dominant and nondominant family size to be relevant in this language. Given that morphological families in Dutch are both smaller and semantically more cohesive, we argued that this result supported the hypothesis that the family-size effect crucially depends on semantic similarity. The introduction of the dichotomy between dominated and nondominated family members is a useful operationalization to take semantic relatedness into account in our analyses. We leave it to future research to develop more fine-grained operationalizations of semantic relatedness within morphological families.

Bates et al. (2003) studied response latencies in picture naming across a broad range of languages. They observed that picture naming latencies in one language could be predicted from the frequency and word-length counts in another language. They interpreted these results as arguing in favor of a substantial semantic component to the word-frequency effect. Following the line of research developed by Bates et al. for the cross-linguistic predictivity of frequency in picture naming and the cross-linguistic predictivity of frequency and family size in Moscoso del Prado Martin et al. (2003), we investigated the cross-language predictivity of frequency and family size across Finnish, Dutch, and Hebrew. We observed substantial cross-language predictivity for frequency across the three languages and more limited cross-language predictivity for word length. This suggests that there is considerable similarity in concept frequency in these languages and that Zipf’s (1949) observation that more frequent words tend to be shorter holds to some extent even across unrelated languages. Following Bates et al., we interpret these results as another indication of a substantial semantic component to the word-frequency effect.

The most important cross-linguistic finding, however, is that the cross-language predictivity of family size is absent when the distance between the morphological systems, as reflected in the degree of stem productivity, becomes very large. Finnish and Hebrew, the languages with the greatest and the smallest stem productivity, respectively, showed no additional predictivity for family size once the within-language measures (frequency, length, and family size) had been taken into account. This lack of predictivity contrasts markedly with the significant predictivity of family size for word length. This suggests to us that there is a higher degree of overlap between the semantic organization in the mental lexicon of morphologically related words in Hebrew and Dutch and in Finnish and Dutch than there is for Finnish and Hebrew. Given that we have thus far only investigated three language families, and only very few languages within these families, this line of explanation remains necessarily tentative and requires further research.

Although the cross-language predictivity of family size suggests that there may be considerable overlap in semantic organization, in the sense that words in dense morphological neighborhoods tend to have translation equivalents that also have dense morphological neighborhoods, the absence of such predictivity for Finnish and Hebrew suggests that there are limits to this cross-language predictivity. To understand why these limits might arise, consider, for instance, the consequences of the different degrees of productivity of compounding in Finnish, Dutch, and Hebrew. In Finnish, compounding is extremely productive; in Dutch, it is productive; and in Hebrew, it is marginally productive at best. Thus, complex concepts expressed by compounds in Finnish will have lexical (instead of phrasal) counterparts in Dutch relatively often but very seldom in Hebrew. In Hebrew, many Finnish words will require phrasal translations. Consequently, the patterns of lexical coactivation in Finnish will resemble the coactivation patterns of their translation equivalents to a much larger degree in Dutch than in Hebrew. If, as has been argued by De Jong et al. (2003), the coactivation of the morphological family members indeed codes the semantic percept of a word, then the present results support the Whorfian view of language, according to which language codetermines thought (see, e.g., Boroditsky, 2001). For languages with similar morphologies, the morphology might guide thought along similar paths, thereby giving rise to considerable cross-language predictivity of family size. When morphological systems are very different, as for Hebrew and Finnish, the well-worn paths along which morphology might lead thought become notably different, as witnessed by the breakdown of the cross-linguistic predictivity of family size for these languages.

References


### Appendix

The Morphological Family of Kirja [Book] in Finnish

<table>
<thead>
<tr>
<th>kirja</th>
<th>book</th>
</tr>
</thead>
<tbody>
<tr>
<td>väräskirja</td>
<td>dissertation</td>
</tr>
<tr>
<td>muistikirja</td>
<td>notebook</td>
</tr>
<tr>
<td>päiväkirja</td>
<td>diary, notebook</td>
</tr>
<tr>
<td>romaanikirjallisuus</td>
<td>novel literature</td>
</tr>
<tr>
<td>aikakauskirja</td>
<td>journal</td>
</tr>
<tr>
<td>kirjasto</td>
<td>library</td>
</tr>
<tr>
<td>lainakirjasto</td>
<td>public library</td>
</tr>
<tr>
<td>kirjastonhoitaja</td>
<td>librarian</td>
</tr>
<tr>
<td>kirjoitus</td>
<td>writing</td>
</tr>
<tr>
<td>kirjoitusjärjestelmä</td>
<td>writing system</td>
</tr>
<tr>
<td>kirjepaino</td>
<td>paper weight</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>kirjailija</th>
<th>author</th>
</tr>
</thead>
<tbody>
<tr>
<td>kirjailijantoiminta</td>
<td>authorship</td>
</tr>
<tr>
<td>asiakirja</td>
<td>document</td>
</tr>
<tr>
<td>kirjailla</td>
<td>embroider</td>
</tr>
<tr>
<td>kirjoittaa</td>
<td>write</td>
</tr>
<tr>
<td>kirja</td>
<td>letter (a written communication)</td>
</tr>
<tr>
<td>kirjain</td>
<td>letter (the symbol)</td>
</tr>
<tr>
<td>kirjeenkantaja</td>
<td>postman</td>
</tr>
<tr>
<td>kirjeenvaaltaja</td>
<td>correspondent</td>
</tr>
<tr>
<td>kirjeenvaalitoveri</td>
<td>pen pal</td>
</tr>
<tr>
<td>kirjoittautaa</td>
<td>register</td>
</tr>
<tr>
<td>kirjoituskone</td>
<td>typewriter</td>
</tr>
</tbody>
</table>

Received July 30, 2003
Revision received January 13, 2004
Accepted January 16, 2004