Bilingualism: Language and Cognition 6 (3), 2003, 213–225 © 2003 Cambridge University Press DOI: 10.1017/S1366728903001147

How word frequency affects morphological processing in monolinguals and bilinguals*

MINNA LEHTONEN
MATTI LAINE
Åbo Akademi University

The present study investigated processing of morphologically complex words in three different frequency ranges in monolingual Finnish speakers and Finnish-Swedish bilinguals. By employing a visual lexical decision task, we found a differential pattern of results in monolinguals vs. bilinguals. Monolingual Finns seemed to process low frequency and medium frequency inflected Finnish nouns mostly by morpheme-based recognition but high frequency inflected nouns through full-form representations. In contrast, bilinguals demonstrated a processing delay for all inflections throughout the whole frequency range, suggesting decomposition for all inflected targets. This may reflect different amounts of exposure to the word forms in the two groups. Inflected word forms that are encountered very frequently will acquire full-form representations, which saves processing time. However, with the lower rates of exposure, which characterize bilingual individuals, full-form representations do not start to develop.

Introduction

Frequency, the number of times a word occurs in a language, is amongst the main factors affecting the speed of word recognition. High frequency words are generally processed faster than low frequency words, and this has been found to be the case both with morphologically simple and complex words (e.g., Gardner, Rothkopf, Lapan and Lafferty, 1987; Taft, 1979). Frequency effects arise in several tasks, such as in reading aloud, picture naming, and semantic and lexical decision.

Frequency of occurrence is also highly relevant to access and representation of morphologically complex word forms, such as *walk* + *ed* or *pill* + *s*. It is commonly assumed that we have two alternative ways to recognize such word forms: either they are accessed via their constituent morphemes (so-called decomposition route, which is slower and more error-prone but spares storage space in long-term memory) or via representations which correspond to whole word forms (so-called full-form or direct route, which is faster but requires more storage space). Irregular forms, such as *drank*, need to be accessed via the full-form route irrespective of their frequency, but for regular inflections, it is possible that frequency is an important determinant of the recognition route used. Indeed, Pinker (1991) argued that high frequency regular forms may be coded into the long-term memory as whole units. Support for the claim has been found in aphasic patients. Laine, Niemi, Koivuselkä-Sallinen and Hyönä (1995) studied the reading performance of a Finnish agrammatic aphasic and found that even though the patient made significantly more errors with low and medium frequency inflected nouns than with corresponding monomorphemic control nouns, the difference vanished when he read highly familiar inflected word forms (e.g., *aamu* + *lla* = ‘morning’ + an adessive case ending, ‘in the morning’). The authors interpreted this as evidence for full-form representations for very high frequency inflected words and decomposed representations for all other inflected forms. A similar finding in an Italian-speaking aphasic was recently reported by Luzzatti, Mondini and Semenza (2001).

With regard to normals, Alegre and Gordon (1999) studied English lexical decision performance in English-speaking individuals and found that the absolute frequency of word forms affects the processing route employed. Their experimental logic was based on systematic manipulation of two frequency factors, CUMULATIVE STEM FREQUENCY and SURFACE FREQUENCY of words. Of particular interest was the point at which surface frequency started to affect reaction times. The frequency manipulation method was first introduced by Taft (1979). Cumulative stem frequency (base frequency, lemma[1]

---

* This study was financially supported by a grant (#20010) from the Joint Committee of the Nordic Social Science Research Councils (NOS-S) to the second author.

[1] Technically, a lemma is a set of word forms with the same stem. The dictionary headword is one form a lemma can take to represent a ‘word’ in all its inflected forms.

Address for correspondence
Minna Lehtonen, Department of Psychology, Åbo Akademi University, FIN-20500 Turku, Finland
E-mail: *minlehto@abo.fi*

frequency) is a measure of the summative frequency of all the inflectional variants of a single word, while surface frequency refers to the frequency of the specific word form, i.e., how often a word occurs in that particular form. An effect in surface frequency manipulation, i.e., shorter reaction times for the high surface frequency words than for the low surface frequency words (cumulative stem frequency is kept constant) is interpreted as evidence for full-form representations at least for the high frequency words. In base frequency manipulation, in turn, shorter reaction times for the high base frequency words than for the low frequency words are interpreted as decomposition for the inflectional word forms in both target groups. Alegre and Gordon (1999) suggested that full-form representations begin to develop for morphologically complex words already when the surface frequency exceeds six occurrences per million. This estimate may apply also for morphologically limited languages other than English, but it is not clear whether it would generalize to languages that are richer in morphology, since the morphological richness of a language has been suggested to be an important factor affecting the organization of the mental lexicon (Hankamer, 1989). For example, Vannest, Bertram, Järvikivi and Niemi (2002) suggested that due to extensive employment of the decomposition process, speakers of a morphologically rich language might have more developed parsers than users of a language that is morphologically poorer. If this is the case, the threshold of developing full-form representations for inflected words in a highly inflecting language might be significantly higher than that proposed by Alegre and Gordon (1999). Moreover, economy of storage considerations (i.e., keeping vocabulary size within some reasonable limits), prompting the use of the decomposition route, may be more acute in languages with rich inflectional morphology.

Unlike many Indo-European languages, Finnish is a language with a very rich morphology. A Finnish noun can have as many as 2000 different inflectional forms (Karlsson, 1983). For example, the inflected word form *talo + i + ssa + mme + kin* ‘also in our houses’ includes five morphemes. In languages with limited inflectional morphology, for instance English or Swedish, many of these inflectional elements would be replaced by free-standing words, such as prepositions.

Because of the agglutinative[2] nature of Finnish, it would be plausible to assume separate (i.e., decomposed) representations for stems and affixes of inflected word forms rather than have all different inflectional forms of a word stored as full forms in the mental lexicon. In fact, both the studies conducted with Finnish-speaking normals using either the method of visual lexical decision (Niemi, Laine and Tuominen, 1994; Laine, 1996; Laine and Koivisto, 1998; Laine, Vainio and Hyönä, 1999) or eye-movement registration (Hyönä, Laine and Niemi, 1995), and those conducted with aphasic patients (Laine, Niemi, Koivuselkä-Sallinen, Ahlsén and Hyönä, 1994; Laine et al., 1995), seem to point in this direction. The main setup used in these studies has been a contrast between otherwise matched inflected and monomorphemic words. If in a visual lexical decision task it takes longer to respond to the inflected items than to the monomorphemic ones, it is generally concluded that the (slower) decomposition route has been used for lexical access. If no difference in the reaction times is found, it is assumed that the processing of inflected words has taken place via the direct route (i.e., full-form representations), in the same way as for the monomorphemic words.

For the Finnish language, the so-called Stem Allomorph/Inflectional Decomposition model (SAID) has been put forth (Laine et al., 1994; Niemi et al., 1994) to account for the data obtained from the empirical studies listed above. SAID is a dual route model, which claims that one of the two access routes is responsible for word recognition. The model states that inflected Finnish words are processed via the decomposition route. The only exception would be high frequency inflected words (e.g., *sauna* + *ssa* ‘sauna’ + ‘in’), which are assumed to be represented as their full forms.

The present study aims to investigate the effects of frequency on morphological processing in the Finnish language. The argument that very high frequency inflected words would be processed as whole word forms is based on a study of a single aphasic patient only (Laine et al., 1995). To obtain further evidence on the effects of frequency on processing of inflected word forms in Finnish, we studied normal, healthy participants with a visual lexical decision task (Experiment 1). The hypothesis we tested was that in a contrast between inflected and monomorphemic nouns representing a very high frequency range (e.g., average surface frequency > 100 per million), there would be no difference in reaction times between the two word groups, while in contrasts including low or medium frequency nouns a significant processing delay for the inflected items would appear.

In addition, word frequency derived from language corpora, the actual exposure to a language is, naturally enough, a crucial factor. In this respect, bilinguals form a particularly interesting group, as due to usage of two languages they receive much less exposure for each one of them than a monolingual speaker of a language does. Moreover, we noted earlier that a language that employs morphology to a great extent may pose different demands for language users concerning the processing strategies and lexical organization than languages that

---

[2] An agglutinative language is a language in which the words are formed by ‘gluing’ morphemes together. Words are made up of a linear sequence of distinct morphemes and each component of meaning is represented by its own morpheme.

are more limited in morphology, for which simpler mechanisms might be sufficient. A theoretically interesting question is whether a person who has acquired both a morphologically rich and a morphologically limited language in childhood has different, separate strategies for morphological processing in the respective languages, or whether one language has affected the representation and the processing of the other.

The study of morphological processing in bilinguals has received scant interest, even though languages differ greatly from one another in terms of morphology. Portin and Laine (2001) were among the first to study this issue by investigating morphological processing in Finnish-Swedish bilinguals. In contrast to morphologically rich Finnish, Swedish is a Germanic language with a rather limited morphological system.[3] In line with an earlier study by Ahlsén (1994), Portin and Laine (2001) found that monolingual Swedes processed both monomorphemic and inflected Swedish nouns at approximately equal speed. In contrast, Finnish-Swedish bilinguals processed inflected Swedish items significantly slower than monomorphemic ones. Thus, it could be that the richness of morphology that the bilinguals have to deal with in Finnish makes them to transfer their typical Finnish processing strategy, decomposition of inflected words, into Swedish. In addition to studying the frequency effects on morphological processing in Finnish monolinguals, the present study investigated how Finnish-Swedish bilinguals process morphologically complex Finnish words (Experiment 2). Also in the bilinguals, three frequency levels (low, medium and high) were employed.

Balanced bilinguals may use their two languages roughly equally often in everyday life, even though estimates are very difficult to make as the two languages may be used in different contexts and thus employ partly different semantic categories of the mental lexicon. Nevertheless, less exposure would mean that in a bilingual's lexicon, the activation level of lexical representations encountered in a language would not rise to the same degree as in a monolingual's mind. In Experiment 2, we employed the same frequency range paradigm in bilinguals as in monolinguals in Experiment 1 and investigated if the frequency effects in morphological processing differed from those obtained in monolinguals. Specifically, we hypothesized that even though monolinguals would show evidence for full-form representations of high frequency inflected forms, bilinguals might still resort to morphological decomposition of the same forms. This was because we assumed that not even the high frequency inflected forms were encountered often enough by a bilingual to develop full-form representations. Our hypothesis, then, was that the bilinguals would show evidence for decomposition with inflected Finnish nouns (i.e., slower reaction times than with corresponding monomorphemic words) throughout the whole frequency range (low, medium and high).

Experiment 1

Experiment 1 was a standard visual lexical decision task with monolingual Finnish participants. The aim was to compare the processing of Finnish inflected vs. monomorphemic nouns over three different frequency ranges. The specific hypothesis derived from the SAID model was that with Finnish-speaking monolinguals, those Finnish inflections that had a low or medium surface and cumulative stem frequency would be processed via the decomposition route. However, since the SAID model predicts that very high frequency inflected words might be processed via the full-form route, no difference in reaction times was expected for the inflected–monomorphemic contrast in the high frequency range.

Method

Materials

For the lexical decision task, altogether six 20-item word lists were collected from the Turun Sanomat lexical database (which includes 22.7 million word tokens) using a computerized search program (Laine and Virtanen, 1999). Three of the lists included only monomorphemic Finnish nouns and the other three only inflected Finnish nouns with 6–8 different case forms per list. Each list of monomorphemic items was contrasted with a list of inflected items. One paired list of monomorphemic vs. inflected nouns was selected from a low (surface) frequency range, one from a medium frequency range and one from a very high frequency range. The contrasted word lists were matched for the following factors: surface frequency, bigram frequency[4] and cumulative stem frequency, word length and morphological family size.[5] Cumulative stem frequency and family size were matched between monomorphemic and inflected words only WITHIN each frequency range. Words with high surface frequency generally tend to have high cumulative stem frequency and family size as well, and this

---

[3] Swedish nouns can be affixally marked for definiteness and they are inflected for gender and number (e.g., *blomma* 'flower'; *blomm* + *or* + *na* = 'flower' + plural marker + marker for definiteness, 'the flowers'). There are only up to 8–10 forms of verbs and nouns in Swedish if passives and genitives are included.

[4] Bigram frequency stands for the average frequency of all two-letter sequences in a word.

[5] Morphological family size refers to the number of derivations and compound words where the stem in question appears. A word with a large family size has been found to be processed faster than a word with a small number of family members (Schreuder and Baayen, 1997; Bertram, Baayen and Schreuder, 2000).

Table 1. *Properties of target words in Experiment 1.*

Word category	WL	SF	BF	BiF	FS	A/C
LOW FREQUENCY						
Monomorphemic	6.60	0.37	1.75	1239	12.4	3.88
Inflected	6.60	0.38	1.73	1251	11.7	3.58
MEDIUM FREQUENCY						
Monomorphemic	6.75	17.8	79.0	1067	302.7	3.39
Inflected	6.75	17.8	79.9	1053	301.7	3.41
HIGH FREQUENCY						
Monomorphemic	6.50	104.9	547.1	1295	1054	2.86
Inflected	6.50	102.1	537.1	1287	1086	2.85

Mean Values of Word Length (WL) in letters, Surface Frequency (SF), Base Frequency (BF), Bigram Frequency (BiF), Family Size (FS) and Abstractness/Concreteness (A/C) for the different word groups. Surface and Base Frequency are reported as frequencies per million.

tendency could not be controlled throughout the conditions. However, considering the influence of these three frequency factors on morphological processing, only the surface frequency of the inflected word could be the plausible factor that might contribute to the development of full-form representation for a particular inflected form. The characteristics of all three pairs of word lists are presented in Table 1. (See the Appendix for all the word lists.)

In addition, the word lists were included in a preliminary questionnaire, in which six participants estimated the abstractness/concreteness levels of each word form using a 5-point scale (1 = very abstract, 5 = very concrete). The contrasted word lists were then compared with each other for their average abstractness/concreteness estimates, and it could be concluded that in all pairs, the word lists had comparable abstractness/concreteness ratings (see Table 1).

A feature of Finnish is that inflection may be accompanied by phonological changes in the stem[6] (e.g., *kukka*: 'a flower' → *kuka* + *n* 'flower's', not **kukka* + *n*; *kastike* 'dressing' → *kastikkee* + *ksi* 'to dressing', not **kastike* + *ksi*). It has been suggested that phonological transparency is among factors that affect word recognition (Frauenfelder and Schreuder, 1992). However, based on lexical decision results with inflectional forms representing different stem changes, Niemi et al. (1994) claim that the different allomorphs of decomposed noun stems in Finnish have separate mental representations (see also Järvikivi and Niemi, 2002). Nevertheless, in this experiment, as many as possible phonologically transparent stem and affix combinations were selected. This meant that the words were mainly selected so that the base form would remain orthographically and phonologically unchanged by inflection. In some cases this was not possible, but the number of items with a stem change was minimized (< 30% in all lists of inflected words). Also, to rule out any effects that might be due to the excellent Swedish knowledge of the bilingual group, no Swedish cognates were selected into the word lists.

In sum, Experiment 1 included three 40-word sets of nouns, yielding altogether 120 target words. 93 filler words (57 Finnish nouns in nominative singular form and 36 nouns in inflected form) were also included.

In addition to the target words and fillers, 213 nonwords (50% of the total number of items) were included in the stimulus set. They were constructed by changing one to three letters of real Finnish words, such that the phonotactic rules of Finnish were not violated. Many of them also had endings similar to the affixes of real words (e.g., *motkussa* could be considered a nonword **motku* + an inessive case ending *-ssa*). The 'case endings' of the nonwords included nominative, partitive, genitive, adessive, inessive, elative and illative endings. The distribution of these endings followed the distribution of the case forms in the set of real words (target and filler items), and thus the participants could not determine the right answer (word/nonword) on the basis of the endings. Finally, in three nonwords, the stem was a real word, but the 'case ending' was a non-existing one. These were included to prompt the participants to read the whole letter string and not only the base to determine the correctness of the word (e.g., *nosturiska* could be analysed as consisting of the real word 'crane' *nosturi* 'crane' + a non-existing ending **-ska*, which is close to real case endings such as *-ssa* and *-sta*).

Participants

Twenty university students participated in Experiment 1 after giving their written informed consent. None of them reported reading difficulties and everyone had normal or

---

[6] Note that Finnish is a fully regular language and thus the phonological changes in the stem are directly reflected in the orthography.

Table 2. *Self-evaluation of language skills in the monolingual participants.*

	Excellent	Good	Satisfactory	Deficient
Self-evaluation of ability to:				
• talk in Finnish in everyday context	95%	5%	–	–
• talk in Swedish in everyday context	–	15%	75%	10%
• read ordinary written Finnish (e.g., newspapers)	95%	5%	–	–
• read ordinary written Swedish (e.g., newspapers)	–	50%	50%	–

corrected-to-normal vision. Ten of the participants were male and ten were female. The age of the participants was 21–27 years (mean, 23.7; SD, 1.13). In the Finnish translation of the Edinburgh Inventory (Oldfield, 1971), all participants reported being completely or mostly right-handed. All participants were native speakers of Finnish and monolingual in the sense that they had acquired only the Finnish language before school age (i.e., before the age of seven) and gone through the country's educational system in Finnish. Finland is officially a bilingual country, and in comprehensive and secondary school it is obligatory to study Swedish. Thus, all of the monolingual participants had also studied Swedish at school for at least 6 years. The participants estimated their language skills both in Finnish and in Swedish using a 4-point scale (1 = deficient, 2 = satisfactory, 3 = good, 4 = excellent). Everyone rated their skills in Finnish to be excellent or good while in Swedish only from deficient to good (mean estimate for Finnish skills, 3.9; SD, 0.3; Swedish skills, 2.28; SD, 0.55). The difference in the self-estimates between the two languages was highly significant ($t(19) = 13.6$, $p < .001$). See Table 2 for the self-evaluation data of these participants.

To further estimate the participants' language skills, a digit reading task with 20 items was included. Chincotta, Hyönä and Underwood (1997) found that digit reading times in bilinguals were faster in their dominant language when effects of different word length of digits in the two languages were controlled. Here the participants read in Finnish numbers from 1 to 20 presented in random order. The reading time of each participant was measured with a stopwatch. The average reading time for these participants was 11.5 seconds (SD, 2.04). We will return to this result when describing Experiment 2.

Procedure

For the visual lexical decision task, the participants were instructed to decide as quickly and accurately as possible whether the letter string appearing on the computer screen was a real Finnish word or not and to respond by pressing a response key with their dominant hand. By pressing the left button (with index finger) the participant would report seeing a real word and by pressing the right button (with middle finger) he/she would report seeing a nonword. Each participant completed the handedness evaluation scale before the task proper. At the end, the participants filled the language skills questionnaire.

The setup for the visual lexical decision task was as follows: an asterisk appeared on the screen, and the participants were to direct their gaze onto it. After 500 milliseconds, a string of letters appeared to the right of the previously shown asterisk. After giving their response (yes to a word, no to a nonword), the participants pushed a third button (with the non-dominant hand) at their own pace to make the next stimulus appear on the screen. The experiment was run by a specially made computer program (SuperLab Experimental Laboratory Software, Version 2.0, Cedrus Corporation) which recorded the participants' reaction times in milliseconds (the time from the appearance of the letter string to the pressing of the reaction time key) and the correctness of their responses. The task was presented in two blocks. The blocks were constructed so that the number of items from each item category (monomorphemic, inflected, low, medium and high frequency, fillers and nonwords) was equivalent. Every participant performed both of the blocks, with a five-minute break in-between, during which he or she filled a background information questionnaire. The presentation order of the blocks was counterbalanced, i.e., half of the participants performed the task with Block 1 first while the other half did Block 2 first. The presentation order of the words (both real and nonwords) within each block was randomized for each participant. Before the experiment proper, participants had a practice session of 30 items to familiarize themselves with the task. The words and nonwords of this set were not the items used in the real experiment. Participants were tested individually in a quiet room. The experiment took about 45 minutes. All communication during the experiment was in Finnish.

Results

Prior to data analyses, incorrect responses and response latencies longer than three standard deviations above individual mean value were discarded. None of the participants exceeded the preset error rate criterion of 15%. The error rates of the participants varied between

Table 3. *Mean Response Latencies (in milliseconds) and Error Rates with Standard Deviations (by Participants) for the stimulus types in Experiment 1.*

Word category		RT (SD)	Error (SD)
Low frequency	Monomorphemic	639 (94)	0.05 (0.04)
	Inflected	742 (177)	0.14 (0.08)
Medium frequency	Monomorphemic	567 (101)	0.01 (0.02)
	Inflected	587 (101)	0.02 (0.03)
High frequency	Monomorphemic	567 (146)	0.02 (0.03)
	Inflected	580 (113)	0.01 (0.02)

1.67% and 5.73% (mean, 3.44; SD, 1.13). No item was discarded from the analysis as all of them were classified correctly by over half of the participants. The average by-participant reaction times and error percentages per condition can be found in Table 3.

Two-way ANOVAs (frequency range × morphological structure) were performed for RTs and error rates. The analysis of variance for reaction times revealed significant main effects both in the by-participant and the by-item analysis for frequency (F1(2, 38) = 89.1, $p < .001$; F2(2, 114) = 74.7, $p < .001$) and for morphological structure (F1(1, 19) = 68.6, $p < .001$; F2(1, 114) = 26.6, $p < .001$). These main effects indicated that high frequency words were generally processed faster than low frequency ones and also morphological structure played a role in reaction times by slowing RTs for inflected items. There was also a significant interaction between the two factors in both the by-participant and the by-item analysis (F1(2, 38) = 5.9, $p < .05$; F2(2, 114) = 10.9, $p < .001$), confirming that the processing of inflected vs. monomorphemic words differed between frequency levels. The pattern of the by-participant results is depicted in Figure 1.

The ANOVA for the errors yielded a significant main effect for frequency both in the by-participant and the by-item analysis (F1(2, 38) = 16.3, $p < .001$; F2(2, 114) = 24.0, $p < .001$). Also a significant main effect for morphological structure was found in the by-participant and the by-item analysis (F1(1, 19) = 25.0, $p < .001$; F2(1, 114) = 7.6, $p < .01$). These main effects stem from the fact that low frequency words elicited more errors than high frequency ones, and so did the inflected items in contrast to monomorphemic words, at least in the low frequency range. The interaction term in the error rate analysis was also significant (F1(2, 38) = 16.3, $p < .001$; F2(2, 114) = 6.8, $p < .01$), reflecting the fact that the difference in error rates between monomorphemic and inflected items was far greater in the low frequency condition than in the medium or high frequency condition.

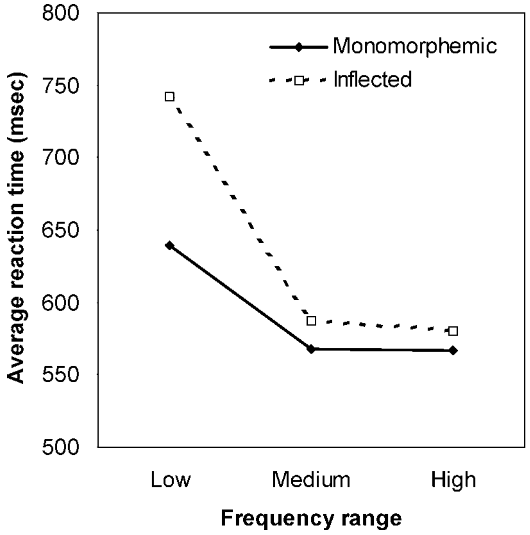

Figure 1. Experiment 1: Mean lexical decision latencies for the monolingual participants.

For all the pairwise inflected–monomorphemic reaction time contrasts, a two-tailed paired t-test for participants and a two-tailed standard two-sample t-test for items was performed. In the low frequency range, inflected items elicited significantly slower reaction times than monomorphemic ones both in the by-participant analysis ($t1(19) = 4.41$, $p < .001$) and in the by-item analysis ($t2(38) = 5.11$, $p < .001$). Inflected items also received more errors than the monomorphemic items, and the difference was significant both in the by-participant and in the by-item analysis ($t1(19) = 5.23$, $p < .001$; $t2(25,8) = 2.80$, $p = .01$).

As regards the medium frequency contrast, inflected words elicited again longer reaction times than the monomorphemic ones. The difference was significant in the by-participant analysis ($t1(19) = 3.27$, $p < .01$), but did not reach significance in the by-item analysis ($t2(38) = 1.82$, $p = .076$). The difference in error rates was not significant in the by-participant ($t1(19) = 1.83$, $p > .05$) or in the by-item analysis ($t2(26, 5) = 1.80$, $p = .05$).

In the high frequency contrast, the difference in reaction times was not significant in the by-participant ($t1(19) = 0.71$, $p = .489$) or in the by-item analysis ($t2(38) = 1.03$, $p = .308$). Also in the error rates, there was no significant difference between item types in the by-participant ($t1(19) = 1.17$, $p = .258$) or in the by-item analysis ($t2(29, 4) = 1.17$, $p = .251$).

Discussion

We observed a significant difference in reaction times and error rates between monomorphemic and inflected words in the low frequency list and a significant difference in reaction times in the medium frequency list, suggesting morphological decomposition for those inflections. This has been a robust finding for most inflected nouns also in earlier studies of Finnish. However, a new finding was the fact that the significant difference in reaction times to monomorphemic words disappeared when the frequency of morphologically complex words was high, which points to full-form processing for these inflections. A shift from decomposed to full-form processing may, however, be gradual. It is plausible that already some of the medium frequency inflected words had developed full-form representations since the difference in that condition manifested itself only in the by-participant analysis, not for items or error rates. Familiar inflected word forms thus seem to be processed like monomorphemic ones, via the full-form route, and the prediction of the SAID model received support.

The reaction times for monomorphemic words in the medium frequency and high frequency range were equally long (567 ms) in this experiment (see Table 3). Thus, one could speculate that the lack of difference in the high frequency range was due to a 'floor' effect: the participants could not have physically responded any faster even though they had actually RECOGNIZED the monomorphemic words faster than the inflected words in that frequency range. If this were the case, one would assume the individual standard deviations per condition of reaction times to decrease when moving from medium frequency range to the high one, i.e., when getting closer to the 'floor' level. However, this does not happen (mean SD, calculated from individual participants' RTs, for medium frequency: monomorphemic, 84; inflected, 93; mean SD for high frequency: monomorphemic, 97; inflected, 97), and therefore it is plausible to assume that the participants indeed recognized both the monomorphemic and inflected words of the high frequency range at equal speed.

Experiment 2

In Experiment 1 we found evidence suggesting that monolingual Finnish speakers decompose low frequency inflected nouns, but for high frequency inflected words, the direct route seems to be employed. This change may be gradual, as already some of the medium frequency inflected words seem to have developed full-form representations. For the bilinguals, it would be interesting to see whether a partly different pattern of results would appear. Specifically, the hypothesis tested in Experiment 2 was that with the bilingual group, in all frequency ranges there would be a processing delay suggestive of morphological decomposition for the inflected items when contrasted with monomorphemic control words.

Method

Materials and procedure

The materials and the experimental procedure in this experiment were the same as the ones used in Experiment 1.

Participants

Nineteen university students participated in Experiment 2 after giving their written informed consent. None of the participants reported reading difficulties, and all had normal or corrected-to-normal vision. Four of the participants were male while 15 were female. The age of the participants was 19–31 years (mean, 25.2; SD, 3.31). In the Edinburgh Inventory (Oldfield, 1971) 18 participants reported being completely or mostly right-handed, and one participant was left-handed. All participants were early Finnish-Swedish bilinguals, and had learned to speak both languages before school age (i.e., before the age of seven). Most of them (84%) had one or both parents speaking Finnish to them, so it is likely that the exposure to Finnish had indeed begun at a very early age. All of the participants had attended school with Swedish as the language of instruction but used also Finnish with friends and in the society (see Table 4). They estimated their language skills both in Finnish and in Swedish using a 4-point scale (1 = deficient, 2 = satisfactory, 3 = good, 4 = excellent). As most of them had acquired the two languages from the start and were recruited by the criterion of being bilingual, they probably used similar norms for estimating skills in both of their languages (i.e., estimates for two mother tongues, not for a mother tongue and a strong second language). All of them rated themselves to be excellent or good in both languages (mean estimate for Finnish skills, 3.79; SD, 0.41; Swedish skills, 3.89; SD, 0.31). There was no significant difference in the self-estimated skills between the two languages ($t(18) = 1.00$, $p = .21$), and they did not differ significantly from the group of Experiment 1 in their Finnish skill estimates either ($t(37) = 1.06$, $p = .298$). Thus it can be concluded that the participants were quite balanced bilinguals, and their everyday language skills in Finnish were approximately similar to those of the monolingual group. As expected, the bilingual group differed significantly from the monolingual group of Experiment 1 as regards their Swedish skills ($t(31{,}5) = 13.89$, $p < .001$).

The digit reading task for the bilinguals was the same as for the monolinguals in Experiment 1. The average reading time in Finnish for this group was 11.7 seconds (SD, 2.00), and it did not differ significantly from that of the monolingual group ($t(37) = .303$, $p = .764$). This is further evidence for the claim that the Finnish skills of the bilinguals were at the same level as those of the monolinguals.

Results

The statistical analyses were performed in the same way as in Experiment 1. No data of the participants had to be rejected using the error rate criterion of 15%, but two items from the low frequency group of inflected words were discarded from the analyses because they had received

Table 4. *Language background data and self-evaluation of language skills in the bilingual participants.*

LANGUAGE BACKGROUND				
Languages used during childhood	Mother: Swedish Father: Finnish	Mother: Finnish Father: Swedish	Both parents: Finnish, Swedish outside home	Both parents:Swedish, Finnish outside home
	26%	32%	26%	16%
	Swedish	Finnish	Both	Other
Language at primary and secondary school	95%	–	5%	–
Language at high school	89%	–	–	11%
Language at university	84%	5%	11%	–
	Swedish only	Finnish only	Both	
Current language use at home, university and work	5%	–	95%	
SELF-EVALUATION OF THE ABILITY TO SPEAK AND READ SWEDISH AND FINNISH				
	Excellent	Good	Satisfactory	Deficient
Self-evaluation of ability to:				
• talk in Finnish in everyday context	79%	21%	–	–
• talk in Swedish in everyday context	84%	16%	–	–
• read ordinary written Finnish (e.g., newspapers)	79%	21%	–	–
• read ordinary written Swedish (e.g., newspapers)	95%	5%	–	–

an incorrect response from over 50% of the participants. The individual error rates varied between 0.72% and 9.63% (mean, 5.00; SD, 2.36), and the mean error rate for these participants was significantly higher than that of the monolingual group ($t1(37) = 2.70$, $p = .01$). Also the average by-participant reaction times for each condition were significantly slower in the bilingual group when compared to those of the monolingual group.[7] The average by-participant reaction times and error rates can be seen in Table 5.

Table 5. *Mean Response Latencies (in milliseconds) and Error Rates with Standard Deviations (by Participants) for the Stimulus Types in Experiment 2.*

Word category		RT (SD)	Error (SD)
Low frequency	Monomorphemic	791 (104)	0.08 (0.07)
	Inflected	900 (144)	0.14 (0.11)
Medium frequency	Monomorphemic	673 (97)	0.00 (0.01)
	Inflected	724 (112)	0.01 (0.02)
High frequency	Monomorphemic	641 (89)	0.01 (0.02)
	Inflected	714 (126)	0.02 (0.02)

The two-way ANOVA for reaction times revealed a significant main effect both in the by-participant and the by-item analysis for frequency ($F1(2, 36) = 94.2$, $p < .001$; $F2(2, 112) = 55.3$, $p < .001$) and for morphological structure ($F1(1, 18) = 45.9$, $p < .001$; $F2(1, 112) = 29.3$, $p < .001$). This indicates that high frequency items were generally processed faster than low frequency words and that inflected items elicited overall longer reaction times than monomorphemic ones. The interaction term failed to reach significance ($F1(2, 36) = 2.98$, $p > .05$; $F2(2, 112) = 2.06$, $p = .132$), suggesting that processing of inflected vs. monomorphemic words was approximately similar in all frequency levels with the bilingual participants. Their RT pattern can be seen in Figure 2.

The ANOVA for error rates yielded a significant main effect both in the by-participant and in the by-item analysis for frequency ($F1(2, 36) = 37.2$, $p < .001$;

---

[7] The only exception were the monomorphemic items in the high frequency range where the difference was only nearly significant ($p = .063$).

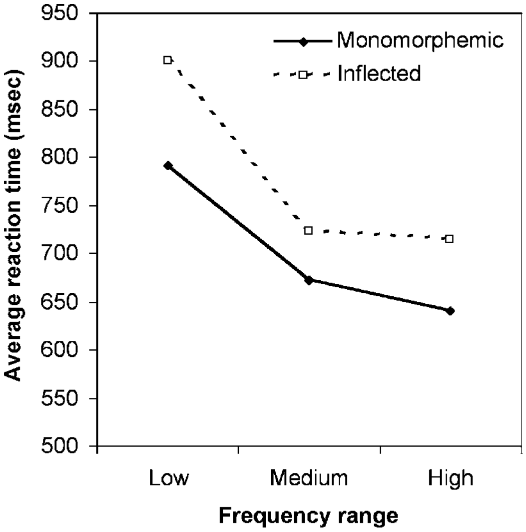

Figure 2. Experiment 2: Mean lexical decision latencies for the bilingual participants.

F2(2, 112) = 25.5, $p < .001$) and for morphological structure (F1(1, 18) = 11.5, $p < .01$; F2(1, 112) = 4.73, $p < .05$). This is due to the greater number of errors for the low frequency items than the higher frequency ones and also more errors for the inflected items than for the monomorphemic words. The interaction term reached significance in the by-participant analysis but not in the by-item analysis (F1(2, 36) = 5.49, $p < .05$; F2(2, 112) = 2.13, $p = .124$). The interaction indicates that error rates tended to differ with morphological structure in different frequency ranges.

For all the contrasts, a two-tailed paired t-test for participants and a two-tailed standard two-sample t-test for items was performed. In the low frequency range, the inflected items elicited significantly slower reaction times than the monomorphemic ones both in the by-participant analysis ($t1(18) = 4.55$, $p < .001$) and in the by-item analysis ($t2(36) = 3.13$, $p < .01$). The low frequency inflected items also received more errors than the monomorphemic control items, but the difference was significant only in the by-participant analysis ($t1(18) = 2.90$, $p < .01$), not in the by-item analysis ($t2(36) = 1.72$, $p > .05$).

With respect to the medium frequency items, inflected targets were again significantly slower than the monomorphemic ones both in the by-participant and in the by-item analysis ($t1(18) = 4.71$, $p < .001$; $t2(38) = 2.59$, $p < .05$). There was no significant difference in error rates either in the by-participant ($t1(18) = 1.71$, $p = .104$) or in the by-item analysis ($t2(25{,}1) = 1.51$, $p = .144$).

In the high frequency range, there was still a significant difference in reaction times between inflected and monomorphemic items both in the by-participant ($t1(18) = 4.34$, $p < .001$) and in the by-item analysis ($t2(38) = 4.93$, $p < .001$). Again, there was no significant difference in error rates between inflected and monomorphemic items either in the by-participant ($t1(18) = 1.14$, $p = .268$) or in the by-item analysis ($t2(38) = 0.89$, $p = .378$).

Due to different effects obtained in Experiments 1 and 2 for the high frequency lists, a two-way ANOVA (morphological structure × language group) was performed for RTs and error rates of those items. Main effects for both morphological structure (F1(1, 37) = 11.55, $p < .01$; F2(1, 76) = 18.12, $p < .001$) and language group (F1(1, 37) = 8.22, $p < .01$; F2(1, 76) = 99.95, $p < .001$) were found both in the by-participant and in the by-item analysis, suggesting that the bilinguals processed the inflected items more slowly than the monomorphemic ones and that the bilingual group was slower overall. Also, the interaction term was significant (F1(1, 37) = 5.52, $p < .05$; F2(1, 76) = 7.93, $p < .01$), stemming from the fact that only the bilingual group exhibited a difference between monomorphemic and inflected targets in the high frequency range. In the error rates, no main effects arose for morphological structure (F1(1, 37) = .036, $p = .850$; F2(1, 76) = .043, $p = .835$) or language group (F1(1, 37) = .018, $p = .893$; F2(1, 76) = .043, $p = .835$), and the interaction term was also non-significant both in the by-participant and in the by-item analysis (F1(1, 37) = 2.61, $p = .115$; F2(1, 76) = 2.13, $p = .149$).

Discussion

The results of the bilingual group showed that inflected nouns were processed significantly more slowly than monomorphemic nouns in all three frequency ranges. In contrast to monolinguals, even high frequency items yielded a significant processing delay for inflections, suggesting that the bilinguals employed morpheme-based access even with the most familiar inflected words used in the experiment. This supports the hypothesis that bilinguals have not been exposed to these words often enough to have developed whole-word representations for them. Also the fact that the bilingual group was slower overall in contrast to monolinguals suggested that the subjective frequencies of words were lower for them.

General discussion

We set out to test the hypothesis that word frequency affects the visual recognition of morphologically complex Finnish nouns. The participants were presented with words from three frequency ranges, high, medium and low, and morphological processing was studied in Finnish-speaking monolinguals and Finnish-Swedish bilinguals. The prediction was that the monolinguals would have full-form representations for high frequency inflected words and decomposed for all other inflected forms. With regard to bilinguals, we hypothesized that they would employ the decomposition route for inflections at all frequency levels, including the highest one. The results of the study supported these predictions: in Experiment 1 (with monolingual participants) a significant difference in reaction times between monomorphemic and inflected items vanished in the high frequency range, whereas in Experiment 2 (with bilingual participants), the difference between the two word types remained significant in all three frequency ranges. This suggests that with monolingual Finnish speakers, some inflected nouns can be processed via the full-form route as the SAID model (Niemi et al., 1994) and Pinker (1991) propose. However, according to our results, processing of inflected Finnish nouns via the full-form route would seem to be a rare or non-existing phenomenon for bilinguals.

A central explanation for the difference between the groups is presumably the amount of exposure to the words. As inflected word forms are encountered more and more often, it will become economical to represent them as a whole since that guarantees fast access to them. Therefore, monolingual Finns employ full-form processing with high frequency inflected words even though they process medium (at least to some extent) and low frequency inflected words via the decomposition route. Bilinguals, on the other hand, get much less lexical input in a particular language than monolingual speakers of that language do. In this study, this is reflected also in the overall difficulty of the task: bilinguals were slower in the task overall. As a result, full-form representations for inflected word forms would not develop.

Another potential factor affecting the difference between the bilinguals and the monolinguals in the high frequency range could be that due to the lower exposure to the language, they were more insecure about the correctness of inflected forms. Thus, they might have conducted a post-lexical check making sure that the particular stem and affix really fit together, thereby making their responses slower. If this were the case and the word types had been otherwise equally easy to process for them, a more careful examination of these stimulus words should lead to speed–accuracy trade-off, i.e., to lower error rates than for the less-carefully examined monomorphemic words. However, as both the reaction times and the error rates are higher for the inflected items, the processing delay cannot be caused by a speed–accuracy trade-off.

The high frequency words in this study had a surface frequency of over a hundred per million while for the medium frequency words the same measure was under 20. It is probable that with monolinguals, already some of the medium frequency inflected words were processed as a whole, since the reaction time difference between word types in that group was markedly reduced and not any more significant in the item analysis when compared with the low frequency items. Anyhow, it seems that somewhere near these numbers inflected Finnish word forms start having full-form representations in monolingual speakers. These numbers are far greater than the limit of full-form representations (six occurrences per million) suggested by Alegre and Gordon (1999). One explanation for a higher limit in our study could be that economy of processing constraint plays a role here. In other words, development of full-form representations for inflectional forms with a frequency over six per million might expand the mental lexicon too much in a morphologically rich language like Finnish.

A future study investigating frequency effects in morphological processing will be one that looks into the processing of inflected nouns in a morphologically limited language like Swedish throughout different frequency ranges, in the same way as with the Finnish language in this study. Using such a setup, we could see whether Swedish speakers really employ full-form route also with the very low frequency inflected words, and whether even bilinguals possibly start developing full-form representations at some point of the frequency continuum.

As mentioned in the introduction to this paper, an interesting question in bilingualism is whether bilingual speakers of two typologically very different languages recognize morphologically complex words in a different way from monolingual speakers of these languages and whether one language affects the processing of the other. Portin and Laine (2001) found that with Swedish inflections, bilinguals employed a different strategy (decomposition) than monolinguals who seemed to process inflected nouns as a whole. In the present study, we found that bilinguals used the decomposition route also with Finnish nouns. This is otherwise in line with the results obtained from monolingual Finnish speakers except for the high frequency items. Thus, it seems that bilinguals have a similar style of morphological processing in Finnish as monolinguals. Having native-like command also of a structurally distant language, Swedish, does not apparently directly affect the processing of Finnish. The exposure to the Swedish language only makes the encounters with Finnish words less common, and thus full-form representations cannot develop for these inflections, as they do in monolinguals. If cross-language morphological transfer exists, it is probable that it is Finnish that influences morphological processing in Swedish and not the other way around because morphological factors are so essential in the Finnish language. This asymmetry of effects that languages seem to have to one another could be further tested by studying morphological processing in language pairs that have limited morphological systems, for example, in the Swedish-English bilingual context. Then we could see if there are any differences in bilinguals between processing of these two languages and whether any asymmetries would appear.

The method employed in this study, contrasting monomorphemic words with otherwise similar inflected words, has been commonly used in Finnish studies. Bertram (2000), however, criticizes the setup for a frequency matching problem it might provoke. In his hypothetical example, he assumes that even when apparently matched by surface and cumulative stem frequency, inflected vs. monomorphemic target nouns may have developed different resting state activation levels. If the target forms in question are actually recognized via full-form representations (i.e., the inflectional target has occurred often enough to develop a full-form representation which its counterpart, the monomorphemic target, has by definition), inflectional variants of the targets included in the cumulative stem frequency count boost the resting state activation levels of the targets differentially. All decomposed inflectional variants of the monomorphemic target boost the resting state activation of the target form which is the stem (possible stem changes are not taken into account here). In contrast, the inflected target form gets no benefit from the other inflectional variants as their forms do not match. As a result, the inflected target elicits longer reaction times because of its inherently lower resting state activation level in the mental lexicon. While many assumptions are made by Bertram (2000), it suffices to say that his speculation actually predicts a disadvantage for higher frequency inflected targets because those represent the ranges where the inflected targets may develop full-form representations. As is evident from the above, our results in monolinguals point in the opposite direction.

Another problem with the design used is that it does not provide a means for testing a possibility that decomposition always occurs. In this hypothetical situation, a process of re-composition (checking whether the stem and affix fit together) would happen after analysing the morphological components of the word, and the speed of re-composition would vary depending on, for example, word frequency. If all inflected words were decomposed but familiar words were re-composed faster than infrequent ones, the difference in reaction times in contrast to monomorphemic words would decrease in the high frequency group, maybe even disappear altogether. For the bilinguals, the lower amount of exposure would have an effect on the speed of re-composition and thus make the reaction time difference visible even in the high frequency range. However, the aphasia data by Laine et al. (1995), which was not based on reaction time data but on reading errors, suggests that full-form representations do exist for inflected nouns. The patient was able to read highly familiar inflected nouns correctly but had trouble with low and medium frequency ones, suggesting that the most frequent inflected words were processed in the same way as monomorphemic words (with which the patient had much less problems), as whole entities. Note that this was not a speeded task. In addition, there is evidence that during early acquisition of Finnish morphology, a child may make errors which indicate that he or she has memorized inflected forms as holistic, unanalysable units (Niemi and Niemi, 1987). In fact, Niemi and Niemi (1987) reported that their participant tended to add a partitive ending into word forms that were already in partitive form. Later on, the child began to take these holistic inflected words more analytically, but it is improbable that the whole word representations formed at an early age would have vanished totally as a result of realizing what morphemes the words contain. These studies provide evidence that even polymorphemic words can have full-form representations in Finnish.

To summarize, we found that frequency affects the recognition of visually presented inflected Finnish nouns. In monolinguals, high frequency and even some medium frequency words can be processed as full forms while others are decomposed, but in Finnish-Swedish bilinguals, only the decompositional representations are normally possible. We suggest that the reason for this is the different amount of exposure to the words in different groups. Inflected word forms that are encountered very frequently will be represented as a whole since that saves processing time. However, if word forms are encountered more rarely, as is often the case in bilinguals, full-form representations do not begin to develop.

References

Ahlsén, E. (1994). Cognitive morphology in Swedish: Studies with normals and aphasics. *Nordic Journal of Linguistics*, 17, 61–73.

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

Bertram, R. (2000). *Morphology in the Mind: Word Processing in Finnish and Dutch*. PhD Dissertation, University of Turku. [Published as Turun yliopiston julkaisuja. *Annales universitatis turkuensis*. Series B, part 235. Humaniora. Painosalama Oy.]

Bertram, R., Baayen, R. H. & Schreuder, R. (2000). Effects of family size for derived and inflected words. *Journal of Memory and Language*, 42, 390–405.

Chincotta, D., Hyönä, J. & Underwood, G. (1997). Eye fixations, speech rate and bilingual digit span: Numeral reading indexes fluency not word length. *Acta Psychologica*, 97, 253–275.

Frauenfelder, U. & Schreuder, R. (1992). Constraining psycholinguistic models of morphological processing and representation: The role of productivity. In G. Booij & J. van Merle (eds.), *Yearbook of morphology 1991*, pp. 165–183. Dordrecht: Kluwer.

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

Hankamer, J. (1989). Morphological parsing and the lexicon. In W. D. Marslen-Wilson (ed.). *Lexical representation and process*, pp. 392–408. Cambridge, MA: MIT Press.

Hyönä, J., Laine, M. & Niemi, J. (1995). Effects of word's morphological complexity on readers' eye fixation patterns. In J. M. Findlay, R. Kentridge & R. Walker (eds.), *Eye movement research: Mechanisms, Processes and Applications*, pp. 445–452. Amsterdam: Elsevier.

Järvikivi, J. & Niemi, J. (2002). Form-based representation in the mental lexicon: Priming (with) bound stem allomorphs in Finnish. *Brain & Language*, 81, 412–423.

Karlsson, F. (1983). *Suomen kielen äänne-ja muotorakenne* [The phonological and morphological structure of Finnish]. Juva. Werner Söderström.

Laine, M. (1996). Lexical status of inflectional and derivational suffixes: Evidence from Finnish. *Scandinavian Journal of Psychology*, 37, 238–248.

Laine, M. & Koivisto, M. (1998). Lexical access to inflected words as measured by lateralized visual lexical decision. *Psychological Research*, 61, 220–229.

Laine, M., Niemi, J., Koivuselkä-Sallinen, P., Ahlsén, E. & Hyönä, J. (1994). A neurolinguistic analysis of morphological deficits in a Finnish-Swedish bilingual aphasic. *Clinical Linguistics and Phonetics*, 8, 177–200.

Laine, M., Niemi, J., Koivuselkä-Sallinen, P. & Hyönä, J. (1995). Morphological processing of polymorphemic nouns in a highly inflecting language. *Cognitive Neuropsychology*, 12, 457–502.

Laine, M., Vainio, S. & Hyönä, J. (1999). Lexical access routes to nouns in a morphologically rich language. *Journal of Memory and Language*, 40, 109–135.

Laine, M. & Virtanen, P. (1999). *WordMill Lexical Search Program.* Centre for Cognitive Neuroscience, University of Turku.

Luzzatti, C., Mondini, S. & Semenza, C. (2001). Lexical representation and processing of morphologically complex words: Evidence from the reading performance of an Italian agrammatic patient. *Brain and Language*, 79, 345–359.

Niemi, J., Laine, M. & Tuominen, J. (1994). Cognitive morphology in Finnish: Foundations of a new model. *Language and Cognitive Processes*, 3, 423–446.

Niemi, J. & Niemi, S. (1987). Acquisition of inflectional marking: A case study of Finnish. *Nordic Journal of Linguistics*, 10, 59–89.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, 9, 97–113.

Pinker, S. (1991). Rules of language. *Science*, 253, 530–535.

Portin, M. & Laine, M. (2001). Processing cost associated with inflectional morphology in bilingual speakers. *Bilingualism: Language and Cognition*, 4, 55–62.

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

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

Vannest, J., Bertram, R., Järvikivi, J. & Niemi, J. (2002). Counterintuitive cross-linguistic differences: More morphological computation in English than in Finnish. *Journal of Psycholinguistic Research*, 31, 83–106.

Received December 30, 2002
Revision received June 26, 2003
Accepted July 2, 2003

Appendix: *Target words*

LOW FREQUENCY

Monomorphemic	Inflected
HAIKARA 'stork'	AIROA 'scull' + partitive marker
JUORU 'gossip'	AMMEEN 'tub' + genitive marker
KARITSA 'lamb'	APILAN 'clover' + genitive marker
KARPALO 'cranberry'	ENOSTA 'uncle' + elative marker
KULKUNEN 'bell'	ERHEEN 'mistake' + genitive marker
LANTTU 'rutabaga'	HAIMASSA 'pancreas' + 'in'
LIESKA 'flame'	KELMULLA 'foil' + adessive marker
MAININKI 'roller'	KIRSTUUN 'trunk' + 'into'
MAJAVA 'beaver'	KITAAN 'maw' + 'into'
NIKSI 'trick'	LEILISSÄ 'skin bottle' + 'in'
OIKKU 'whimsy'	LORUN 'rigmarole' + genitive marker
PINAATTI 'spinach'	LUURIIN 'handset' + 'into'
PIUHA 'string'	NUTTUUN 'jacket' + 'into'
PUSAKKA 'windcheater'	RUSETIN 'bow' + genitive marker
RINKKA 'backpack'	SUPPILON 'funnel' + genitive marker
VANUKAS 'pudding'	TIKKIÄ 'stitch' + partitive marker
VEITIKKA 'rascal'	VAINUN 'scent' + genitive marker

VIINERI ‘Danish pastry’	VEKARAN ‘kid’ + genitive marker
VOKAALI ‘vowel’	VILTIN ‘blanket’ + genitive marker
VONKALE ‘large fish’	VIPUUN ‘lever’ + ‘into’

MEDIUM FREQUENCY

Monomorphemic	Inflected
EMÄNTÄ ‘hostess’	AURINGON ‘sun’ + genitive marker
KAPPALE ‘paragraph’	HAAVAA ‘wound’ + partitive marker
KATAJA ‘juniper’	HERKKUA ‘delicacy’ + partitive marker
KAVERI ‘buddy’	HUIPULLE ‘top’ + allative marker
KOHTALO ‘fate’	HUONEEN ‘room’ + genitive marker
KOKOELMA ‘collection’	ISKUSSA ‘hit’ + ‘in’
KORTTI ‘card’	KIMPPUUN ‘bouquet’ + ‘into’
KURSSI ‘course’	KOIRAA ‘dog’ + partitive marker
KYNNYS ‘doorstep	KUORON ‘choir’ + genitive marker
LENKKI ‘link’	LEHMÄN ‘cow’ + genitive marker
LUETTELO ‘directory’	LUKKOON ‘lock’ + ‘into’
MOOTTORI ‘engine’	MAITOA ‘milk’ + partitive marker
PAKKANEN ‘frost’	MYRSKYN ‘storm’ + genitive marker
PATSAS ‘statue’	POLVEN ‘knee’ + genitive marker
PORUKKA ‘crowd’	REITTIÄ ‘route’ + partitive marker
POTILAS ‘patient’	RUUMIIN ‘corpse’ + genitive marker
SIHTEERI ‘secretary’	TAITOA ‘skill’ + partitive marker
TEHDAS ‘factory’	VALOSSA ‘light’ + ‘in’
URAKKA ‘job’	VARJOON ‘shadow’ + ‘into’
VAHINKO ‘accident’	VELKAA ‘debt’ + partitive marker

HIGH FREQUENCY

Monomorphemic	Inflected
HENKILÖ ‘person’	AAMULLA ‘morning’ + adessive marker
JÄSEN ‘member’	ALUSSA ‘beginning’ + ‘in’
JOUKKO ‘bunch’/‘crowd’	ASIAA ‘issue’ + partitive marker
KANTA ‘base’	AUTOLLA ‘car’ + adessive marker
KAUPPA ‘shop’	HINTAA ‘price’ + partitive marker
KAUPUNKI ‘city’	ILLALLA ‘evening’ + adessive marker
LIIKENNE ‘traffic’	JALKAAN ‘foot ’ + ‘into’
LIITTO ‘union’	KAHVIA ‘coffee’ + partitive marker
LUOKKA ‘class’	KEVÄÄN ‘spring’ + genitive marker
MERKKI ‘sign’	KUNTOON ‘shape’ + ‘into’
MUSIIKKI ‘music’	MATKALLA ‘trip’ + adessive marker
NAINEN ‘woman’	OSALLA ‘part’ + adessive marker
OHJELMA ‘program’	PELIIN ‘game’ + ‘into’
ONGELMA ‘problem’	PUUSTA ‘wood’ + elative marker
PAIKKA ‘place’	RAHAA ‘money’ + partitive marker
POLIISI ‘police’	SYKSYLLÄ ‘autumn’ + adessive marker
TILANNE ‘situation’	SYYSTÄ ‘reason’ + elative marker
TORSTAI ‘Thursday’	VAARASSA ‘danger’ + ‘in’
VALTIO ‘state’	VAUHTIA ‘speed’ + partitive marker
VOITTO ‘victory’	VIIKON ‘week’ + genitive marker