Affixal salience and the processing of derivational morphology: The role of suffix allomorphy

Juhani Järvikivi and Raymond Bertram
University of Turku, Finland

Jussi Niemi
University of Joensuu, Finland

This study explores the relevance of suffix allomorphy for processing complex words. The question is whether structural invariance of the morphological category (i.e., lack of allomorphy) would affect the processing of Finnish derived words. A series of four visual lexical decision experiments in which alternatively surface and base frequency was manipulated showed that the two invariant suffixes, namely denominal –stO and deadjectival –hkO, showed reliable effects of base frequency, whereas for the two categories with suffix allomorphy, deverbal –Us and deadjectival –(U)Us, only surface frequency played a role. A further experiment showed that even with the most frequent variant of –(U)Us, namely –Ude-, response latencies were a function of surface frequency only. It is shown that neither the results from the experiments here nor previous findings from processing Finnish words can be accounted for by suffix frequency, the frequency ratio between the derived word and its base, or morphological productivity in any straightforward manner. We conclude that the lack of allomorphy, that is, structural invariance, significantly adds to affixal salience and therefore enhances morphological decomposition. The implications of this finding for models of lexical processing are discussed.
According to Laudanna and Burani (1995) the processing of derived morphologically complex words is importantly regulated by the likelihood with which a given affix occurs as a processing unit in a language, the so-called affixal salience. Affixal salience pertains to the probability with which an affix is likely to emerge from the orthographic/phonological string as an active unit of processing serving lexical access for morphologically complex words. Thus, the extent to which the affix is salient to the processing system is likely to influence the balance of storage and computation for derived words, in other words, the (un)involvement of morphological structure in lexical processing. Computation simply refers here to the situation where the word-internal morphological structure is actively involved in the real time processing of morphologically complex words without further ontological commitment to the so-called symbolic paradigm. A variety of factors – such as, productivity, frequency, (orthographic) length, and affix confusability – have been shown to affect the relative salience of affixes (e.g., Baayen, 1994; Laudanna & Burani, 1995). In the present study we investigated the possibility that suffix allomorphy has consequences for affixal salience. We will show that whether a particular derivational category is realised by several surface allomorphs or not has processing consequences for Finnish derived words. More precisely, we will show that extensive allomorphy diminishes affixal salience and detracts from morphological decomposition regardless of factors such as affix frequency and productivity. In the following section we will discuss the role of some of the factors affecting affixal salience in more detail. First we will take up the issue of productivity.

**FACTORS AFFECTING AFFIXAL SALIENCE**

Although the importance of morphological productivity has been acknowledged in both linguistics and psycholinguistics, only a handful of processing studies have dealt with this issue directly. Anshen and Aronoff (1988) showed that productivity affected the production of English derived words. They found that productive affixes, such as *-ness*, resulted in a wider variety of word types and a greater number of new words than less productive affixes, such as *-ity*. Thus, the words produced by the less productive affixes were more likely to be listed in a dictionary than the words produced by the more productive ones. Also Baayen (1994) showed that the number of neologisms produced by Dutch speakers was predictable from the productivity of the affix in question: the more productive the process the more previously unattested word types were observed. Laudanna and Burani (1995) presented evidence that the degree of productivity affects the processing of Italian suffixed non-words. The
strings with productive affixes were more likely candidates to undergo morphological parsing than the strings with unproductive affixes. Laine (1996) provided similar evidence for Finnish affixed non-words with both inflectional and productive derivational suffixes. The findings so far suggest differential processing between words with (relatively) low or non-productive and words with (sufficiently) productive affixes. There is no evidence, however, showing that the degree of productivity of affixes has straightforward processing consequences for the recognition of derived words. It seems rather that a certain degree of productivity is needed in order for computational effects to arise (see Vannest, Bertram, Järvikivi, & Niemi, 2002), and productivity – as other affix specific characteristics – is likely to be tied together with a number of other factors, such as (orthographic) length, frequency, and ambiguity.

Laudanna and Burani (1995) showed that affix length and both type and token frequency of the affix affect whether derivational affixes are used in the course of processing in Italian (Laudanna & Burani, 1995, for a review; Burani & Laudanna, 1992; Burani, Dovetto, Thornton, & Laudanna, 1997; Burani, Thornton, Iacobini, & Laudanna, 1995). More precisely, long and frequent affixes were more likely to act as processing units than short and infrequent affixes. Laudanna and Burani (1995) also showed that the degree of confusability – the ratio of an orthographic/phonological string serving as a real affix or pseudo-affix – influences the processing of Italian derivational affixes. It was shown, for example, that non-words containing a prefix ri– with a high real prefix ratio (81 %) were significantly slower to reject than matched non-words with co– with a low real prefix ratio (18 % of occurrences). Analogically, Schreuder and Baayen (1994) showed that obligatory prefix-stripping, postulated by Taft and Forster (1976) for English, would lead to a high proportion of false segmentations in Dutch. A further factor adding to confusability is homonymy. Bertram et al. (2000c) argued that affixal homonymy of the Dutch –er, serving both as an agentive marker ‘werker’ (worker) and a comparative suffix ‘mooier’ (more beautiful), induce holistic processing for words in these suffixes.

Both pseudo-affixation and affixal homonymy represent a deviation from the maximally salient one-to-one correspondence between the affix form and its function in either leading the processor astray where there is no true morphological relation at all, or creating a one-to-many situation with more than one meaning competing for a single form. An opposite type of deviation is created by affix allomorphy, namely that many forms are competing for one meaning.

Schreuder and Baayen (1995) hypothesised that affix allomorphy may be relevant for the processing and representation of polymorphemic words. More precisely, they claimed that the discovery of an affix with several allomorphs, such as the Dutch diminutive suffix with five allomorphs,
namely –je, –tje, –etje, –kje, –pje, would take longer than the discovery of an affix with a single phonological form, such as –heid (–ness). Thus, several forms linked to one meaning may make the discovery of a stable form-meaning relationship more difficult. There is evidence indicating that in lexical acquisition children do have difficulties with the many-to-one mapping of form and meaning (see e.g., Clark’s principles of contrast and simplicity of form; Clark, 1993).

In the present study it will be shown that neither suffix frequency nor productivity alone increase the salience of Finnish derivational suffixes enough for decomposition to take place. Instead, our results suggest that formal invariance – i.e., lack of allomorphy – is needed to enhance the relative affixal salience, so that an affix will serve as a processing unit in lexical comprehension.

EVIDENCE FROM FINNISH DERIVATIONAL MORPHOLOGY

Conceptions from the morphological-typological make-up of morphologically rich languages would lead us to expect that especially the computationally fairly easily segmentable types, such as Turkish and Finnish (see Sproat, 1992), would be the prime candidates for morphological decomposition in lexical comprehension for the huge number of potential word forms alone (Frauenfelder & Schreuder, 1992; Hankamer, 1989; Karlsson & Koskenniemi, 1985). Whereas the evidence suggests that Finnish inflected words tend to be morphologically decomposed (Laine, Vainio, & Hyöna, 1999; Niemi, Laine, & Tuominen, 1994), the recent empirical evidence from Finnish derived words speaks for the opposite. Niemi et al. (1994) proposed that Finnish derived words are recognised holistically in normal comprehension. For supporting evidence, Laine, Niemi, Koivuselkä-Sallinen, and Hyöna (1995) found that the performance of an aphasic deep dyslexic patient was significantly better for derived than for inflected words both in single word-reading tasks and in a repetition experiment. Thus, inflected words were found to induce more errors than derived words in both tasks. The authors also found no statistical difference between derived and comparable monomorphemic words. Hyöna, Laine, and Niemi (1995) studied the processing of Finnish polymorphemic words in a series of experiments where the participants’ eye movements were registered while they performed both visual lexical decision and naming tasks. The authors found no difference between the

1 To be precise, typologically Finnish morphology is fusional-agglutinative in distinction to that of Turkish, which gives us a schoolbook example of agglutinative morphology.
words in the very productive agent marker \(-jA^2\) and the monomorphemic control condition in any of the measures used in the three tasks. In a similar vein, Bertram, Laine, and Karvinen (1999) provided evidence that Finnish derived words formed by \(-jA\), are recognised by whole word representations only. The authors argued that this is due to its being homonymous with an inflectional partitive suffix \(-j+A^3\). Furthermore, they found no difference between words in a marginally productive denominal suffix \(-lA\), e.g., ‘kahvila’ (coffee shop) and monomorphemic words, but did find faster reaction times for words in \(-stO\), a denominal suffix forming collective nouns, e.g., ‘kirja+sto’ (literally: book + collective noun marker \(=>\) library) than for the monomorphemic baseline.

Recently, Vannest, Bertram, Järvikivi, and Niemi (2002) reported a series of visual lexical decision experiments with words formed by denominal adjective forming suffixes \(-isA\), \(-kAs\), and \(-tOn\). \(-isA\) is a mildly productive suffix that forms adjectives denoting abundance of the meaning of the base lexeme (e.g., Penttilä, 1957), e.g., ‘sotaisa’ (warlike, bellicose). \(-kAs\) is used productively to form adjectives of the meaning very BASE-like or BASE-ful, e.g., ‘tehokas’ (effective, efficient). The suffix \(-tOn\), in turn, is a very productive means of forming adjectives meaning BASE-less, e.g., ‘tehoton’ (ineffective, inefficient). In Vannest et al. (2002), the base and surface frequencies of the words formed with each of the three suffixes were systematically varied in order to determine whether they would be accessed via whole word representation or morphemic constituents. For all three suffixes, only whole word frequencies were found to play a role. This was found to be the case despite the fact that all words employed were of relatively low frequency: the average cumulative frequency for the words in \(-isA\), and \(-kAs\) was approximately two occurrences per million, and for the words in \(-tOn\) less than one per million. In contrast, the authors found ample evidence for decompositional access for a number of English derivations. Given the enormous morphological complexity of Finnish – with less than 3 % of normal text being accounted for by monomorphemic nouns and more than 95 % of the morphologically complex word stock occurring less than once per million – the findings strike one as counterintuitive. However, an avenue left unexamined is the possibility that the Finnish suffixes studied

---

2 The capital vowels in \(-jA\) and the suffixes investigated here, namely \(-stO\), \(-hkO\), \(- (U)Us\), and \(-Us\), refer to the archiphoneme marking phonological adjustment following from Finnish vowel harmony. That is, for instance in the case of \(-jA\), the suffix is realised as either /ja/ or /jæ/ depending on whether the stem has a back or front vowel respectively.

3 Although perceptually homonymous with the deverbal \(-jA\), the particular variant of partitive plural inflection is structurally bimorphemic with a combination of a partitive allomorph \(-A\) and a phonologically motivated change of the plural marker \(-i\) to \(-j\) in /V_V/ contexts.
so far have been compromised by factors that yield an affix more or less salient. As discussed above, productivity, frequency, and allomorphy are among these factors. In what follows we take a closer look into the nature of the relevant derivational suffixes.

### FEATURES OF FINNISH DERIVATION

Table 1 lists a selection of Finnish derivational suffixes with cumulative token frequencies, total number of word types, types occurring only once, and indices of productivity in a lexical database compiled from a Finnish newspaper corpus. The index of productivity is based on Baayen’s (e.g., Baayen, 1994) conception of statistical productivity as the likelihood of encountering new unobserved types formed by a given affix in a sample of text. The index of productivity ($P$) for a given affix is calculated by dividing the number of hapaxes ($n_1$) – word types with the given affix occurring only once – by the total number of word tokens with that affix in a corpus of text. The selected affixes are the ones on which empirical data are available (Bertram et al., 1999: denominal –lA, denominal –stO, deverbal –jA).

<table>
<thead>
<tr>
<th>Affix</th>
<th>$N$</th>
<th>$n$</th>
<th>$n_1$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denominal -lA</td>
<td>20249</td>
<td>49</td>
<td>1</td>
<td>0.00005</td>
</tr>
<tr>
<td>Denominal -stO</td>
<td>108561</td>
<td>201</td>
<td>17</td>
<td>0.0002</td>
</tr>
<tr>
<td>Deverbal -jA</td>
<td>311784</td>
<td>2285</td>
<td>681</td>
<td>0.0022</td>
</tr>
<tr>
<td>Denominal -tOn</td>
<td>47497</td>
<td>727</td>
<td>270</td>
<td>0.0057</td>
</tr>
<tr>
<td>Denominal -kAs</td>
<td>52158</td>
<td>157</td>
<td>15</td>
<td>0.0003</td>
</tr>
<tr>
<td>Denominal -isA</td>
<td>13207</td>
<td>59</td>
<td>4</td>
<td>0.0003</td>
</tr>
<tr>
<td>Deadjectival -hkO</td>
<td>2493</td>
<td>147</td>
<td>52</td>
<td>0.0209</td>
</tr>
<tr>
<td>Deadjectival -(U)Us</td>
<td>216499</td>
<td>2703</td>
<td>842</td>
<td>0.0039</td>
</tr>
<tr>
<td>Deverbal -Us</td>
<td>421239</td>
<td>1696</td>
<td>376</td>
<td>0.0009</td>
</tr>
</tbody>
</table>

These types of words are termed hapaxes.

---

4 The frequency and productivity counts are based on the lexical database compiled from seven consecutive annual volumes of a Finnish newspaper Karjalainen (1991–1997). The material consists of 34.5 million word tokens and covers a reasonably long stretch of time to provide a representative view of the use and productivity of derivational affixation. The corpus (Karjalaisen korpus, 34.5 million-word token computer-based newspaper corpus of Finnish based on Karjalainen (Joensuu), compiled by J. Niemi and his associates at the Linguistics Department, University of Joensuu, SGML form created at the Department of General Linguistics, University of Helsinki) is available from Kielipankki at http://www.csc.fi/kielipankki/.
–jA; Vannest et al., 2002: denominal –tOn, denominal –kAs, denominal –isA) and the ones employed in the current study (denominal –stO, deadjectival –hkO, deadjectival –(U)Us, and deverbal –Us).

First, the denominal suffix –lA is represented by only 49 types in the whole of 34.5 million words in the Karjalainen database, with only one single type occurring once. Thus, it is evident that –lA is only marginally – if at all – productive apart from its occasional use in place and proper names (Karlsson, 1983; Penttilä, 1957). Furthermore, both –isA and –kAs seem only marginally productive at best. –isA also has a very productive competitor suffix –inen, forming adjectives with analogical meaning, e.g., ‘sotainen’ (warlike, bellicose). Moreover, the adjectives formed by the two suffixes (–isA and –inen) are formally convergent through part of their inflectional paradigm, thus, in encountering such words as, e.g., ‘sotaisa’ (warlike-partitive-plural) or ‘sotaisilla’ (warlike-adessive-plural), one cannot tell just by looking at those forms which is the relevant base, ‘sotaisa’ or ‘sotainen’. –kAs, in turn, also has a homonymous noun-forming counterpart, which may diminish its affixal salience and enhance its lexicalisation, as was argued to be the case with the deverbal agentive suffix –jA (Bertram et al., 1999). Finally, although the denominal suffix –tOn receives an even higher index of productivity than –jA, it is of lower frequency. As suffix frequency has been shown to contribute to affixal salience (Laudanna & Burani, 1995), it is possible that –tOn is not frequent enough to be used as a processing unit.

In addition to frequency and productivity, however, a further factor potentially influencing the outcome of the previous studies is suffix allomorphy. Table 2 lists partial paradigms for a selection of words formed with Finnish derivational suffixes –stO ‘kirjasto’ (library), –jA ‘jyrsija¨’ (rodent), –tOn ‘arvoton’ (worthless), –(U)Us ‘heikkous’ (weakness), –Us ‘räjähdys’ (explosion), –hkO ‘korkeahko’ (quite high), –kAs ‘raivokas’ (furious), –isA ‘valoisa’ (well-lit, bright), and –lA ‘kahvila’ (coffee shop), respectively. It is immediately evident that all previously studied suffixes, apart from –stO, have more than one suffix allomorph. Furthermore, the low frequent but productive –tOn is realised by two almost suppletive-like allomorphs (–ttOm and –ttOma) throughout most of its inflectional paradigm. Thus it could be that allomorphy may have affected the processing of the words in –tOn resulting in the whole word effects observed in Vannest et al. (2002). The ubiquity of suffix allomorphy in Finnish becomes even more evident if we consider that only five of the 37 Finnish nominal suffixes characterised as rule-based by Karlsson (1983) are formally invariant (i.e., exhibit no suffix allomorphy).

The present experiments investigate the possibility that the suffixes in the previous studies were compromised in terms of affixal salience. More precisely, we will investigate whether they were compromised in affix
**TABLE 2**

Part of the inflectional paradigm for selected words in Denominal –*stO*, Deverbal –*jA*, Denominal –*tOn*, Deadjectival –*(U)Us*, Deverbal –*Us*, Deadjectival –*hkO*, Denominal –*kAs*, Denominal –*isA*, and Denominal –*lA*

<table>
<thead>
<tr>
<th>Case</th>
<th>-stO</th>
<th>-jA</th>
<th>-On</th>
<th>-(U)Us</th>
<th>-Us</th>
<th>-hkO</th>
<th>-kAs</th>
<th>-isA</th>
<th>-lA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nom Sg</td>
<td>kirjasto</td>
<td>jyrsijä</td>
<td>arvoton</td>
<td>heikkous</td>
<td>räjähdys</td>
<td>korkeahko</td>
<td>raivokas</td>
<td>valoisaa</td>
<td>kahvilaa</td>
</tr>
<tr>
<td>Gen Sg</td>
<td>kirjaston</td>
<td>jyrsijän</td>
<td>arvottoman</td>
<td>heikkouden</td>
<td>räjähdyksen</td>
<td>korkeahkon</td>
<td>raivokkaan</td>
<td>valoisan</td>
<td>kahvilan</td>
</tr>
<tr>
<td>Gen Pl</td>
<td>kirjastojen</td>
<td>jyrsijöiden</td>
<td>arvottomien</td>
<td>heikkouksien</td>
<td>räjähdyksien</td>
<td>korkeahkojen</td>
<td>raivokkaiden</td>
<td>valoisi</td>
<td>kahviloiden</td>
</tr>
<tr>
<td>Ptv Sg</td>
<td>kirjastoa</td>
<td>jyrsijää</td>
<td>arvotonta</td>
<td>heikkoutta</td>
<td>räjähdystä</td>
<td>korkeahkoa</td>
<td>raivokasta</td>
<td>valoisaa</td>
<td>kahvilaa</td>
</tr>
<tr>
<td>Ptv Pl</td>
<td>kirjastoja</td>
<td>jyrsijöitä</td>
<td>arvottomia</td>
<td>heikkouksia</td>
<td>räjähdyksiä</td>
<td>korkeahkoja</td>
<td>raivokkaita</td>
<td>valoisia</td>
<td>kahviloita</td>
</tr>
<tr>
<td>Ess Sg</td>
<td>kirjastona</td>
<td>jyrsjänä</td>
<td>arvottoman</td>
<td>heikkoutena</td>
<td>räjähdyksenä</td>
<td>korkeahkona</td>
<td>raivokkaana</td>
<td>valoisan</td>
<td>kahvilana</td>
</tr>
</tbody>
</table>
frequency, productivity, and suffix allomorphy. In order to do this, first, in Experiment 1 we will inspect the role of suffix frequency by using a highly frequent deverbal noun-forming suffix –Us. Second, in Experiment 2, the role of productivity will be investigated with a highly productive deadjectival suffix –(U)Us, the most productive means to form adjective-based nouns in Finnish. Although in fact, –(U)Us attaches to both adjectives and, much less frequently, to nouns, the reason for referring to it as deadjectival in the present study is that only adjective-based nouns formed by –(U)Us are used. Thus, Experiments 1 and 2 investigate whether respectively high frequency and high productivity alone will increase affixal salience enough to induce decomposition for these suffixes. Experiments 3 and 4 investigate whether suffix allomorphy affects affixal salience in Finnish. Experiment 3 investigates the moderative adjective suffix –hkO. –hkO attaches to practically any adjective, for example ‘korkea’ (high) – ‘korkeahko’ (quite high), the only restriction being that the adjective is relational (Karlsson, 1983: 255). Of all derivational suffixes, –hkO is most likely the best candidate to serve as a processing unit in lexical comprehension: it is highly productive, unambiguous (i.e., has no homonymous counterparts), it is semantically predictable, and – as can be seen in Table 2 – it is also formally invariant throughout the inflectional paradigm. Therefore we expect that, alongside the traditional evidence for holistic access to Finnish derivation, words in –hkO will provide evidence for morpheme-based access as well. Finally, in Experiment 4, another invariant suffix, namely the collective denominal suffix –stO, is used to investigate whether the lack of allomorphy still renders a suffix salient enough to function as a processing unit at a lower level of productivity.

**EXPERIMENT 1**

*Words in deverbal –Us.* –Us attaches to verbal bases forming action nouns denoting the act or the result of the verb (cf. English –ing, ‘kloonaus’ [cloning]). It is productive on linguistic accounts (Karlsson, 1983; Penttilä, 1957) and it is extremely frequent in tokens and types as well as quite productive (see Table 1).

In the present study we will use the so-called frequency manipulation technique for all of the experiments. The manipulation is twofold: for one set of items the average surface frequency is varied, while keeping the frequency of the base constant. For another set of items, the base frequency is varied, while the surface frequency is kept constant. A

---

5 This is also reflected in the fact that in the Karjalainen database –(U)Us is attested in noun bases with about 120 types and 12 hapaxes, resulting in a comparably low index of productivity, p = .0001 (cf. Table 1).
positive effect on the former is thought to reflect holistic access of the complex word forms, a positive effect on the latter is interpreted as evidence for morpheme-based access (see e.g., Bertram et al., 2000c; Colé, Beauvillain, & Segui, 1989; Taft, 1979; Vannest et al., 2002). We will come back to the validity of these assumptions after the experiments.

**METHOD**

**Participants.** Twenty students from the University of Joensuu participated in the experiment for a cafeteria coupon (worth about 2 euro). All were native speakers of Finnish and had normal or corrected-to-normal vision.

**Materials.** Twenty words derived by deverbal –Us were selected for both the High Base and Low Base conditions from the Turun Sanomat lexical database (consisting of 22.7 million running words) using the WordMill lexical search program (Laine & Virtanen, 1996). The two sets were matched for surface and bigram frequency and word length as well as base family size (the number of derived and compound descendants of the base form; see, e.g., Bertram, Baayen & Schreuder, 2000a; Schreuder & Baayen, 1997, for the relevance of this factor in lexical processing). By Surface Frequency from here on we mean the cumulative frequency of the derived word itself and its inflectional variants, in other words, lemma frequency. Similarly, Base Frequency refers to the cumulative frequency of the base word and all its inflected forms. In addition, 20 words derived by –Us were selected for the High Surface and Low Surface frequency conditions. Both conditions were matched for base frequency, bigram frequency, and word length as well as base family size. Lexical statistics for each set of words are given in the Appendix, Tables A and B.

Due to an overlap between the items in the base frequency and surface frequency experiments, there were altogether 63 experimental items in deverbal –Us. In addition, the experiment included 131 fillers consisting of monomorphemic (61), derived (32) and inflected (38) words. The experiment also included 194 phonotactically legal non-words with a similar morphological structure as the real words. The non-words were constructed by changing 1–3 letters in an existing Finnish word.

**Procedure.** A visual lexical decision task was run using a Macintosh Power PC running PsyScope 1.2.5 (Cohen, MacWhinney, Flatt, & Provost, 1993). The participants were to decide whether a letter string appearing on the computer screen was a Finnish word or not by pushing the ‘yes’ or ‘no’ buttons on the specially designed button box. Each stimulus was preceded by a fixation point (*) appearing in the middle of the screen for 500 ms; 500
ms after the fixation point disappeared, a stimulus appeared in the middle of the screen in black lowercase Chicago 24-point letters on a light grey background. The items were presented randomised for each participant and time-out was set at 2000 ms from stimulus onset. Twenty practice trials, 10 words and 10 non-words preceded each session. An additional ten trials preceded the first experimental item in each session. All participants were tested one at a time in an experimental room.

**RESULTS**

Before data analyses all incorrect responses as well as responses that were three standard deviations below or above individual means were excluded. The outliers and errors accounted for 5.9% of the data. Table 3 shows the mean response latencies (calculated over correct responses) and error scores (calculated over all responses) for the different experimental conditions. The analyses for the reaction times in the high and low base frequency conditions showed a marginally significant tendency in the participant analysis, $t_1(19) = 2.04, p = .053$. The 16 ms difference between the two conditions proved clearly non-significant in the item analysis, however ($t_2 < 1$). The observed difference in the error rates between the high and low base frequency conditions proved non-significant as well, $t_1(19) = 1.37, p > .2, t_2 < 1$. Expectedly, the high surface frequency condition elicited significantly faster reaction times than the low surface frequency condition, $t_1(19) = 5.95, p < .001, t_2(38) = 3.12, p < .005$. The difference in the error rates was also significant, $t_1(19) = 2.98, p < .01$, $t_2(38) = 2.23, p < .05$.

Overall, the results indicate that the marginal difference in the participant analysis for base frequency is not generalisable. However, to further inspect the possible involvement of the base frequency, a post hoc stepwise regression analysis was carried out with surface and base frequency as well as base family size entered as predictors. The results confirmed the initial expectation that only the whole word frequency emerged as a factor predicting a significant amount of the observed

<table>
<thead>
<tr>
<th>Condition</th>
<th>RT (SD)</th>
<th>Errors (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Base</td>
<td>688 (128)</td>
<td>5.2</td>
</tr>
<tr>
<td>Low Base</td>
<td>704 (128)</td>
<td>6.8</td>
</tr>
<tr>
<td>High Surface</td>
<td>650 (117)</td>
<td>1.8</td>
</tr>
<tr>
<td>Low Surface</td>
<td>705 (135)</td>
<td>6.8</td>
</tr>
</tbody>
</table>
variation in the reaction times, $R^2 = .272$, $F(1, 61) = 22.83$, $p < .001$. Thus, it seems that the words in –Us are commonly processed through their whole word form only. This is a rather surprising result, given the high overall frequency of the suffix. Although, suffix frequency has been shown to influence lexical processing, it seems that, after all, mere high frequency alone is not sufficient to increase affixal salience for the words in –Us enough for morpheme-based processing to take place. It is possible, then, that despite its frequency, the affixal salience of –Us may be diminished by its not being productive enough for a base frequency effect to appear. We will investigate this possibility in the next experiment.

EXPERIMENT 2

Words in deadjectival –(U)Us. In order to further promote affixal salience and increase our chances of finding a base frequency effect, we selected another suffix for experiment 2, namely deadjectival suffix –(U)Us. –(U)Us is one of the most frequent suffixes and certainly the most productive one to derive nouns from adjective bases in Finnish (Karlsson, 1983). It attaches to practically any adjective producing nouns that denote quality or state of (being) the BASE (cf. English –ness), e.g., ‘ahneus’ (greed). Much less productively, it may also attach to nouns, e.g., ‘miehuus’ (manhood), ‘orjuus’ (slavery) (Vesikansa, 1977). Table 1 also shows that –(U)Us is attested with an extremely high number of types of which almost a third are hapaxes (i.e., occur only once in the database). Thus, it also receives a high index of productivity. In sum, with respect to both productivity and frequency –(U)Us is a good candidate for morpheme-based access.

METHOD

Participants. Twenty-two students from the University of Joensuu participated in the experiment for a cafeteria coupon (worth about 2 euro). None had participated in Experiment 1. All were native speakers of Finnish and had normal or corrected-to-normal vision.

Materials. Forty words derived by –(U)Us, 20 for the High Base and 20 for the Low Base conditions, were selected from the Turun Sanomat lexical database using the Laine and Virtanen (1996) WordMill lexical search program. The two sets were matched for surface and bigram frequency, word length, and base family size. In addition, 20 words were also selected for the High Surface and Low Surface frequency conditions. The words in the two conditions were matched for base and bigram frequency, word length and base family size. Lexical statistics for each set of experimental words can be found in the Appendix, Tables A and B.
Due to an overlap between the items in the base frequency and surface frequency experiments, there were altogether 69 experimental items in –(U)Us. In addition, 131 real word fillers consisting of monomorphemic (61), derived (32), and inflected (38) words were included in the experiment. Also, 200 phonotactically legal non-words with a morphological structure corresponding to the real words used in the experiment were constructed by changing one to three letters from the beginning of existing Finnish words.

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

RESULTS

Before data analyses all incorrect responses as well as responses that were 3 standard deviations below or above individual means were excluded. The outliers and errors accounted for 3.2% of the data. Table 4 shows the mean response latencies (calculated over correct responses) and error scores (calculated over all responses) for the different experimental conditions.

Rather surprisingly, the mean response latencies for the high and low base frequency conditions showed an inverse frequency effect with higher mean response latencies for the High Base condition than the Low Base one. The statistical analyses further showed that the difference was significant in the participant analysis, \( t_1(21) = 2.88, p < .01 \), but failed to reach significance in the item analysis, \( t_2(38) = 1.44, p > .1 \). Also, the difference in the error rates between the two conditions was not statistically significant, \( t_1(21) = 1.31, p > .2, t_2 < 1 \). Expectedly, the words in the high surface frequency condition, in turn, were responded to significantly faster than the words in the low surface frequency condition, one-tailed \( t_1(21) = 1.97, p < .05, t_2(38) = 2.00, p < .05 \). However, the observed difference in the error scores was not statistically significant \( (t_1, t_2 < 1) \).

<table>
<thead>
<tr>
<th>Condition</th>
<th>RT (SD)</th>
<th>Errors (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Base</td>
<td>664 (115)</td>
<td>2.9</td>
</tr>
<tr>
<td>Low Base</td>
<td>641 (101)</td>
<td>1.8</td>
</tr>
<tr>
<td>High Surface</td>
<td>616 (106)</td>
<td>1.8</td>
</tr>
<tr>
<td>Low Surface</td>
<td>641 (115)</td>
<td>2.0</td>
</tr>
</tbody>
</table>
The unexpected direction of the frequency effect found in the base frequency manipulation could be at least partly due to the fact that the average surface frequency was somewhat higher for the low base condition than the high base condition (2.1 and 1.5 per million, respectively). However, to further assess the observed inverse base effect, a multiple stepwise regression analysis for all experimental items was carried out with base and surface frequency as well as base family size as predictors. The results indicate that only the surface frequency predicts a significant amount of the observed variation in the response latencies for the experimental words, $R^2 = .165$, $F(1, 67) = 13.25, p < .005$. A post-hoc stepwise regression analysis for the items employed in the base frequency manipulation alone further confirms this observation, despite the observed tendency in the factorial design, Surface Frequency: $R^2 = .185$, $F(1, 38) = 8.61, p < .01$. Thus, it seems that also words formed by the deadjectival suffix –(U)Us are processed via whole word forms only, despite its high frequency and productivity. This suggests that other factors may be involved in decreasing the affixal salience of the suffixes studied so far.

Various studies have indicated that a lack of consistent one-to-one correspondence between the affix form and its function has consequences for morphological processing. This has been shown to be the case with both affix confusability – the ratio of a form serving as a real or a pseudo-affix (e.g., Laudanna & Burani, 1995) – and affixal homonymy, where a single affix form serves two or more morphological functions (e.g., Bertram, Baayen, Schreuder, Laine, & Hyöna, 2000b). Schreuder and Baayen (1995) note that affix allomorphy, that is, many-to-one correspondence between the affix form and affix meaning, may affect the discovery of a consistent form-meaning relationship, and therefore also have consequences for affixal salience. Looking back at Experiments 1 and 2, it is possible that allomorphy was, among other factors, responsible for diminishing the salience for the suffixes in question. In order to assess the role of suffix allomorphy in rendering an affix more or less likely to serve as a unit in lexical processing, the next experiment attempts to increase the affixal salience further by employing a highly productive, formally invariant, suffix, namely –hkO.

EXPERIMENT 3

Words in deadjectival –hkO. In order to investigate whether affixal salience is increased by structural invariance (i.e., lack of suffix allomorphy), the present experiment uses words in –hkO, a very productive suffix attaching to adjective bases with the meaning quite BASE. With a few exceptions that are more like preferences than constraints, it may attach to any adjective. Table 1 shows that –hkO
receives even a little higher index of productivity than the suffix –U(U)s used in Experiment 2. However, in contrast to the previously employed suffixes, –hkO is invariant in form in all inflectional variants. In other words, it does not have any other allomorphs outside the low-level phonological variants accounting for Finnish vowel harmony. Furthermore, it is also unambiguous in other respects: it has neither competitors nor homonymous suffixes. Thus, it might be that in addition to the traditional effects of whole word processing found so far, –hkO is salient enough for a base frequency effect to appear.

**METHOD**

*Participants.* Twenty students from the University of Joensuu participated in the experiment for a cafeteria coupon (worth about 2 euro). None of the participants had participated in the previous two experiments. All were native speakers of Finnish and had normal or corrected-to-normal vision.

*Materials.* Twenty-one words derived by –hkO were selected for the High Base and Low Base conditions from the Karjalainen lexical database using Laine and Virtanen (1996) WordMill lexical search program. Both sets were matched for word length, surface and bigram frequency as well as base family size.

In addition, 19 words derived by –hkO were selected for the High Surface and Low Surface conditions. The sets were matched for base and bigram frequency, word length, and base family size. Lexical statistics for each set of words are given in the Appendix, Tables A and B.

Due to an overlap between the base frequency and surface frequency items, the experiment included altogether 71 experimental items in –hkO. Moreover, 131 real word fillers consisting of monomorphemic (61), derived (32), and inflected (38) words were added. In addition, 202 phonotactically legal non-words with a morphological structure corresponding to the real words were constructed by changing one to three letters from the beginning of an existing Finnish word.

*Procedure.* The procedure was identical to that in the previous experiments.

**RESULTS AND DISCUSSION**

Before data analyses all incorrect responses as well as responses that were 3 standard deviations below or above individual means were excluded. The outliers and errors accounted for 7.0% of the data. In addition, one item, kallihko (quite expensive), belonging to both the High Base and Low
Surface condition was excluded due to an error rate higher than 35%. The removal of the error-prone item did not affect the matching of the sets in either the Surface or the Base conditions. Table 5 shows the mean response latencies (calculated over correct responses) and error scores (calculated over all responses) for the different experimental conditions. The analysis of the base frequency experiment showed that the words in the High Base condition were responded to significantly faster than words in the Low Base condition, \( t_1(19) = 3.00, p < .005, t_2(39) = 1.89, p < .05 \). The High Base condition also induced fewer errors than the Low Base condition. The difference was significant in the participant analysis but not in the item analysis, \( t_1(19) = 2.46, p < .05, t_2(39) = 1.13, p > .1 \). The high surface frequency condition, in turn, elicited faster reaction times than the low surface frequency condition. The difference was significant in the participant analysis, \( t_1(19) = 2.46, p < .05 \), and close to significant in the item analysis, \( t_2(35) = 1.57, p = .06 \). The difference in error rates was marginally significant in the participant analysis but non-significant in the item analysis, \( t_1(19) = 1.59, p = .06, t_2 < 1 \). It seems, then, that both base and whole word representations play a role in the processing of words in \( -\text{hkO} \).

It was hypothesised above that suffix allomorphy would contribute to affixal salience along with other factors such as frequency and productivity. Experiment 3 with \( -\text{hkO} \), a suffix with no allomorphy, indeed indicated that structural invariance of an affix increases affixal salience, importantly enhancing decomposition for words with the suffix in question. Until the present investigation, the processing evidence for Finnish derivations has been largely a continuing series of whole word frequency effects, even in the face of high productivity categories with low frequency words. There is.

---

<table>
<thead>
<tr>
<th></th>
<th>RT (SD)</th>
<th>Errors (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Base</td>
<td>635 (115)</td>
<td>4.5</td>
</tr>
<tr>
<td>Low Base</td>
<td>667 (134)</td>
<td>6.9</td>
</tr>
<tr>
<td>High Surface</td>
<td>619 (101)</td>
<td>3.4</td>
</tr>
<tr>
<td>Low Surface</td>
<td>643 (129)</td>
<td>6.1</td>
</tr>
</tbody>
</table>

\( ^6 \) The results from another run of the surface frequency experiment with different items (16 per condition) and 25 participants showed a significant difference between the High Surface (674 ms) and Low Surface (722 ms) conditions, \( t_1(24) = 4.05, p < .001, t_2(39) = 2.46, p < .05 \), confirming that surface frequency is indeed a significant factor in the processing of words in \( -\text{hkO} \).
a caveat, however: as was noted earlier, –hkO has an even higher index of productivity (see Table 1) than the highly productive –(U)Us. Although unlikely, it is possible that high suffix productivity alone could have been enough to induce the morpheme-based effects observed for the words with this suffix. As the present argument presupposes a ceteris paribus situation, it is important that we make sure that the results from Experiment 3 were not due to the high productivity of –hkO alone. Therefore, to rule out the potential influence of productivity, in the next experiment another invariant – however only mildly productive – denominal suffix, namely –stO, was used to investigate the issue.

EXPERIMENT 4

Words in denominal –stO. –stO is a moderately productive denominal suffix forming collective nouns, e.g., ‘kirja + sto’ (literally: book + collective noun marker = > library), with a productivity index close to that of –isA and –kAs. Its productivity is also significantly lower than that of both –jA and –tOn, neither of which have shown decompositional effects in the previous studies. More importantly, –stO is also significantly less productive than any of the three other suffixes used in the present study, as well as less frequent than the suffixes in Experiments 1 and 2.

Bertram et al. (1999) found that words derived by –stO were responded to faster than the monomorphemic nouns used as controls in visual lexical decision. They argued that this indicated that the words in –stO benefit from a simultaneous activation of the morpheme-based and whole word processing routes through so-called statistical facilitation. There is a caveat, however, as Bertram (2000: 32–33) noted; the technique of using monomorphemic words as baselines may be fallacious in that it is possible that the observed differences spring from different levels/phases of processing, and further, that matching on the different types of frequencies is problematic, since one needs to make assumptions beforehand about whether complex words are listed holistically or decomposed. Therefore, the current base-surface frequency manipulation technique comparing derived words with derived words is needed to certify whether processing of words in –stO is indeed to some extent morpheme-based.

METHOD

Participants. Twenty students from the University of Joensuu participated in the experiment for a cafeteria coupon (worth about 2 euro). None had participated in the previous experiments. All were native speakers of Finnish and had normal or corrected-to-normal vision.
Materials. Thirty-four words derived by –stO were selected from the Karjalainen lexical database (consisting of 34.5 million running words), 17 for the High Base and 17 for the Low Base conditions, using the Laine and Virtanen (1996) WordMill lexical search program. Both conditions were matched for word length, surface and bigram frequency. It was not possible to match for family size of the base, but post hoc analysis will assess whether family size has an independent role to play in processing derived words in –stO.

In addition, 19 derived words in –stO for the High Surface and Low Surface conditions were selected. These conditions were matched for base and bigram frequency, word length, and family size of the base. Lexical statistics for each set of experimental words are given in the Appendix, Tables A and B.

Due to an overlap between the base frequency and surface frequency items, the experiment included altogether 64 experimental items in –stO. In addition, 116 real word fillers consisting of monomorphemic (47), derived (32) and inflected (37) words were included. Finally, 180 phonotactically legal non-words with a morphological structure corresponding to the real words were constructed by changing one to three letters in an existing Finnish word.

Procedure. The procedure was identical to that of the previous experiments.

RESULTS AND DISCUSSION

Before data analyses all incorrect responses as well as responses that were 3 standard deviations below or above individual means were excluded. The outliers and errors accounted for 6.6% of the data. Table 6 shows the mean response latencies (calculated over correct responses) and error scores (calculated over all responses) for the different experimental conditions. For the base frequency experiment, the reaction time analysis again

<table>
<thead>
<tr>
<th></th>
<th>RT (SD)</th>
<th>Errors (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Base</td>
<td>687 (126)</td>
<td>3.2</td>
</tr>
<tr>
<td>Low Base</td>
<td>738 (119)</td>
<td>6.6</td>
</tr>
<tr>
<td>High Surface</td>
<td>677 (121)</td>
<td>4.4</td>
</tr>
<tr>
<td>Low Surface</td>
<td>709 (115)</td>
<td>9.4</td>
</tr>
</tbody>
</table>
showed that the high frequency condition was processed significantly faster than the low frequency condition, $t_1(19) = 3.54, p < .005; t_2(32) = 2.87, p < .01$. In addition, the low frequency condition resulted in a significantly higher percentage of errors than the high frequency condition, $t_1(19) = 2.99, p < .01; t_2(32) = 2.17, p < .05$. As to the surface frequency experiment, the analyses revealed significantly faster reaction times for the high surface frequency condition than for the low surface frequency condition, $t_1(19) = 4.12, p < .001; t_2(36) = 2.17, p < .05$. The difference in error rates was significant in the analysis by participants, but did not reach significance in the item analysis, $t_1(19) = 2.15, p < .05; t_2(36) = 1.78, p > .08$.

As the matching for the base family size in the two conditions was less than perfect, a stepwise multiple regression analysis was carried out for all experimental items with surface and base frequency and base family size as predictors. This analysis revealed that both surface and base frequency, but not family size, are significant predictors for the observed variance in the response latencies, $R^2 = .352, F(2, 52) = 13.68, p < .001$; Surface Frequency, $\beta = -.414$; Base Frequency, $\beta = -.378$). These results support the study of Bertram et al. (1999) with a different paradigm. Thus, it indeed seems that both morpheme-based and whole word representations play a role in the processing of words in –stO.

Taken together with the earlier data from the processing of Finnish derived words the present experiments indicate that suffix allomorphy decreases affixal salience and detracts from decomposition for the categories with several suffix allomorphs. However, there are two possibilities to interpret the results so far: First, it may be that suffix allomorphy exerts a general difficulty for decomposition process resulting in the whole word effects observed in Experiments 1 and 2 as well as in the earlier studies. Second, however, it is possible that the processing of other, less frequent, allomorphic variants of the suffix in question is suppressed by the most frequently occurring allomorph.

As Table 7 shows, deadjectival suffix –Us is mostly realised with the allomorphic variant –Ude-. Also, the basic form of deverbal –Us has a strong competitor in the allomorphic variant –Ukse- and when allomorph clusters are considered, it can be seen that the Uks-cluster is more common than the –Us-cluster. Thus, for both suffixes in Experiment 1 and 2, at least one allomorphic variant has a higher frequency of occurrence than the ones used in the experiments. In order to assess the potential role of allomorph frequency a further experiment was carried out using inflected words in deadjectival –Us with its most frequent allomorph –Ude-.
EXPERIMENT 5

Inflected words in deadjectival –Ude-. As is shown in Table 7, –Ude- comprises almost 36% of all occurrences of the suffix –Us. –Ude- occurs in various case-inflected forms where it is always followed by an inflectional suffix denoting the category in question. The most common category formed with –Ude- is the genitive singular. The genitive singular of the words in –U(U)s is formed with the suffix allomorph –U(U)de- and a genitive inflectional suffix –n, e.g., ‘heikkouden’ (weakness – genitive-singular), and is morphologically segmented as heikko + ude + n. In order to investigate whether allomorph frequency is the decisive factor underlying the observed whole word effects in Experiments 1 and 2, the present experiment used genitive singular inflected words in deadjectival –Us. If higher allomorph frequency increases the affixal salience for –Ude-, we expect to find a positive base frequency effect for the inflected words in question. If, however, suffix allomorphy exerts a more general difficulty for the decomposition process, no base frequency effect should be observed.

METHOD

Participants. Twenty-two students from the University of Turku were paid to participate in the experiment. None had participated in the previous experiments. All were native speakers of Finnish and had normal or corrected-to-normal vision.

Materials. The 69 experimental items in deadjectival –U(U)s from Experiment 2 were inflected in genitive singular. As in Experiment 2, the

<table>
<thead>
<tr>
<th>Allomorph</th>
<th>Percentage of occurrence</th>
<th>Absolute number of occurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td>da-Us</td>
<td>35.7</td>
<td>77290</td>
</tr>
<tr>
<td>Ude</td>
<td>25.4</td>
<td>54991</td>
</tr>
<tr>
<td>Us</td>
<td>17.7</td>
<td>38320</td>
</tr>
<tr>
<td>Ut</td>
<td>14.0</td>
<td>30310</td>
</tr>
<tr>
<td>Uks</td>
<td>7.3</td>
<td>15804</td>
</tr>
<tr>
<td>dv-Us</td>
<td>40.8</td>
<td>172708</td>
</tr>
<tr>
<td>Us</td>
<td>36.6</td>
<td>154173</td>
</tr>
<tr>
<td>Uks</td>
<td>22.7</td>
<td>95621</td>
</tr>
</tbody>
</table>

Table 7: Percentage and absolute number of occurrence for each specific allomorph of Deadjectival –(U)Us and Deverbal –Us.
items were allocated to a Base Frequency (High or Low) and/or a Surface Frequency condition (High or Low), so that each of the four conditions comprised 20 items. The real word and non-word filler materials were the same that were used in Experiment 2, except that the 69 non-words corresponding to the morphological structure of the experimental words were also inflected in genitive singular.

Procedure. The procedure was identical to that of the previous experiments.

RESULTS AND DISCUSSION

Before data analyses all incorrect responses as well as responses that were 3 standard deviations below or above individual means were excluded. The outliers and errors accounted for 7.3% of the data. Table 8 shows the mean response latencies (calculated over correct responses) and error scores (calculated over all responses) for the different experimental conditions. For the base frequency experiment, the reaction time analysis showed that the high frequency condition was processed slower than the low frequency condition. The difference in reaction times was significant in the subject analysis, $t_1(21) = 3.23, p < .01$, but non-significant in the item analysis, $t_2(38) = 1.65, p > .1$. Also the difference in the error scores was statistically non-significant, $t_1(21) = 1.14, p > .2, t_2(38) < 1$. As to the surface frequency experiment, the analyses revealed significantly faster reaction times for the high surface frequency condition than for the low surface frequency condition, $t_1(21) = 6.10, p < .001; t_2(38) = 3.70, p < .001$. There was no difference in the error rates.

In line with the tradition in terminology, we have also here used the term Surface Frequency to reflect the cumulative frequency of the word and all its inflected forms. It should be noted though that the surface frequency proper (here the genitive form) highly correlates with the surface frequency as we defined it. That is, for the base frequency experiment, the genitive frequencies are matched (High: 0.2; Low: 0.3; $t < 1$), for the surface frequency experiment they are manipulated (High: 4.3; Low: 0.3; $p < .001$).
In order to assess further the reverse base frequency effect as well as the effect of the surface frequency difference between the high and low surface conditions, a series of regression analyses was carried out. Stepwise multiple regression for all experimental items with surface and base frequency and base family size as predictors revealed that only surface frequency appeared as a significant predictor for the observed variance in the reaction times, \( R^2 = .229, F(1, 67) = 19.93, \beta = -.479, p < .001 \). A separate stepwise multiple regression analysis for the base frequency items alone showed that even in this item set surface frequency was the only statistically significant predictor, \( R^2 = .161, F(1, 38) = 7.28, \beta = -.401, p < .01 \).

In sum, the results from Experiment 5 showed that the inflected words in –(U)Ude- were processed via the stem and there was no reliable effect of the base. This suggests that the frequency of a particular suffix allomorph is not alone responsible for the observed whole word effects in the previous experiments. Thus, allomorphy seems to induce a more general cost for the processing system decreasing the affixal salience for the suffixes with several allomorphs. As to how this might take place will be more thoroughly discussed in the general discussion, but first we will turn to some methodological issues and some potentially confounding factors.

**METHODOLOGICAL ISSUES**

Over the past 25 years, the base-surface frequency paradigm has been extensively used to assess whether complex words are subject to morpheme-based and/or holistic access. More recently, scholars have been arguing that a surface frequency effect cannot be automatically taken as evidence for whole word access and the lack of a base frequency effect does not automatically imply that morpheme-based access has not taken place. Taft (2004), for instance, proposes a decomposition account in which all morphologically complex words are in a first phase accessed via the base after which – in a second phase – information of the base and the affix are combined. Suppose plurals like ‘rock+s’ and ‘moon+s’ are matched on the frequency of the base, that is, ‘rock’ and ‘moon’ are of equal frequency. Suppose next that ‘rock+s’ is a more frequently occurring formation, that is, has a higher surface frequency than ‘moons’. It is possible that a positive surface frequency effect between these two types of words does not reflect faster whole word access for ‘rocks’ than for ‘moons’, but instead a faster composition process of stem and suffix in the former case.

Consider now the case in which no base frequency effect is observed (as in the current Experiments 1 and 2), when high base frequency items like ‘moons’ are pitted against low base frequency items like ‘cliffs’. A high base frequency word like ‘moons’ would be processed faster in the first
phase than ‘cliffs’, because the stem ‘moon’ is much more frequent than the stem ‘cliff’. In the second stage, however, combining information of the stem ‘moon’ with information of the suffix –s would be importantly slower than combining information of ‘cliff’ and the plural suffix, since ‘moon’ and –s seldom co-occur whereas ‘cliff’ and –s do. This implies that, despite a lack of base frequency effect, decomposition could still have taken place. In other words, in our conceptual world ‘moon’ is mostly experienced as a singular entity, whereas ‘cliff’ can be pluralised without any problem. Taft claims that many high base frequency items in a base frequency experiment are in some sense ‘moon’-like, since the base-surface ratio is by definition lower than the base-surface ratio for low base frequency items. In other words, a pure decomposition model could still be defended, even though for a certain set of complex words one might find a positive surface frequency effect with no base frequency effect.

In contrast to a pure decomposition model, a dual route model like the one advocated by Baayen, Dijkstra, and Schreuder (1997) and Bertram et al. (2000c) assumes that decompositional access always takes place, but simultaneously with holistic access. In this framework a positive surface frequency effect accompanied by no base frequency effect would be interpreted as the result of a relatively slow decomposition process and a relatively fast holistic process. According to Schreuder and Baayen (1995), a number of factors may affect the speed of the decomposition route. Subsequent studies by Baayen et al. (1997) and Bertram et al. (1999, 2000b; 2000c) presented empirical evidence confirming the impact of some of these factors. More specifically, they showed that base frequency effects come and go depending on factors such as affix productivity, affixal homonymy and word formation type. The accompanying claim was that when factors like these are unfavourable for decomposition, complex words are effectively processed in a holistic manner. Evidence for this kind of architecture is not only found in Dutch and Finnish. In a recent study, New, Brysbaert, Segui, Ferrand, and Rastle (2004) argue that a dual route model could accommodate their surface and base frequency effects on English and French familiar plural nouns much better than either pure whole word or pure decomposition models. However, in line with Taft’s framework, it could perhaps still be claimed that the processing is less smooth for high base frequency words than for low base frequency words in the second phase of decomposition. The question then is whether the outcome of base and surface frequency experiments generate any insight in complex word processing at all. It is rather demoralising to think that the

---

8 Although they do note that the connectionist three-layer network model of Davis, van Casteren, and Marslen-Wilson (2003) can also accommodate their pattern of findings.
interpretation of the effects are strongly linked to the theoretical accounts one wishes to defend, or, to turn it around, that the theoretical implications of the outcome of the experiments are not as straightforward as was originally thought. However, what can be upheld is that a positive base frequency effect reflects smooth decomposition or at least effective use of the base morpheme in the course of processing. In contrast, when no base frequency effect occurs, one would have to conclude that for some reason the processing system could not take advantage of a high-frequency base. In general terms, one can claim that the decomposition procedure is troubled. It is either not working at all, it is working very slowly from the start, or alternatively, it starts out working smoothly, but runs into trouble in a later phase.

How could we then interpret the data from a series of experiments as presented here? As we described, every single affix is different in many dimensions, perhaps even more than we can think of. This is why we have chosen to look at one single suffix at a time. At the same time, we have attempted to keep certain aspects as constant as possible across the experiments. Thus, for all suffixes but one (–stO) in this study and the Vannest et al. (2002) study the high base–low base ratio is around 15, the high surface–low surface ratio is about 20, word length is between 7 and 9 characters and the non-manipulated frequency measures are kept as constant as possible. Whenever this was not possible, regression analyses of the whole set of data were carried out for additional insights. Even then, it is possible that various potential confounds were overlooked. We will come back to this in the next section. However, using the base/surface frequency manipulation paradigm and creating as similar circumstances as possible for distinct suffixes creates a situation in which one can assess whether the quantitative and/or qualitative properties of a suffix enhance decomposition. In practice this means that, if for one suffix a base frequency effect can be found and for another suffix – manipulated in a similar way – not, the difference can be ascribed to the properties of the suffixes under investigation. In our case, it turned out to be decisive whether a suffix had several allomorphs or not.

**POTENTIALLY CONFOUNDING FACTORS**

It is important to establish that the base frequency effects from –stO and –hkO stem from the formal invariance of their inflectional paradigm and not from any other factors. In the following we will consider three potential confounds, namely, affix frequency, the base-surface frequency ratio, and productivity. In principle, each of these factors could have been more favourable for a base frequency effect to occur in case of –stO and/or –hkO than in –Us and –(U)Us.
First, Laudanna and Burani (1995) showed that affix frequency is an important factor considering the likelihood of an affix to serve as a unit in lexical processing. We can derive from Table 1 that this variable can be dismissed relatively easily. Only a whole word frequency effect was found for the deverbal –Us despite its much higher overall frequency than that of –stO or –hkO. The situation is not altered if we consider only the occurrences in nominative singular for –Us and the inflections without phonological modifications (99,000 altogether). Even this subgroup of –Us is more frequent than all instances of –stO and –hkO. Overall this applies to the deadjectival –(U)Us as well. The situation is even clearer if we compare the results from Experiment 2 with –hkO with the results from the denominal –tOn in Vannest et al. (2002): –tOn is by far the more frequent of the two as is shown in Table 1. Since in all cases, especially with –(U)Us, type frequency also works for our argument, we can conclude that the present results are not an artifact of the relative suffix frequency of the experimental words.

Second, another important factor is the ratio between the surface and base frequencies of the words employed. Cole, Segui, and Taft (1997) argued that the balance of storage and computation is affected by the relative frequency of the whole word form in contrast with the frequencies of its components. They found effects of morpheme-based recognition only when the morpheme frequency (i.e., the frequency of the base of the complex word) was higher than the frequency of the whole word. Vice versa, holistic effects were observed only in cases where the complex word was more frequent than its base. Recently, Hay (2001) has provided further evidence suggesting that the frequency ratio of the base and whole word form is an important factor in the processing of derived words, especially in the case of low frequency words. Hay (2001) showed that derived words that are more frequent than their bases are perceived as being morphologically less complex than words with higher base than whole word frequency. For the contrasted pairs we selected here and in Vannest et al. (2002), the base frequency was almost always higher than the surface frequency (except for the low base frequency condition of deverbal –Us). In other words, if morpheme-based access had been enhanced by a base-surface ratio higher than 1, we should have found base frequency effects for most suffixes.

Third, the issue of affix productivity should be taken up. Table 1 lists the so-called hapax-conditioned corpus-based count of productivity for all categories under discussion. The rationale for the measure is that it expresses the expectancy rate for new (unobserved) types when N tokens have been sampled (Baayen & Lieber, 1991). That is, affixes with a high number of low frequency types are likely to receive a high index of productivity, the implication being that they are also less likely to be
lexicalised than affixes with higher token frequency per type. In that respect, the adjective suffix –hkO with the highest index of productivity is indeed favoured for decompositional processing. However, it may be argued that in fact, globally, both the denominal –tOn and deadjectival –(U)Us are even more productive: they have an extremely high number of low frequency occurrences, attested by the number of hapaxes, and a very high productivity ratio. Both are also regarded as attachable to almost any base of the relevant category, and have (almost) no restrictions on their scope of application. As it is usually thought that, overall, productivity involves potentiality in both respects (Aronoff, 1976; Baayen & Lieber, 1991; Dressler, 1998), we may claim that both –tOn and –(U)Us should benefit from productivity as much as –hkO. More importantly, however, the denominal –stO, investigated in Experiment 4, is, apart from –lA, the least productive suffix we have employed. In contrast to –stO, however, no base frequency effect could be obtained for any of the more productive suffixes except for the invariant –hkO.

Finally, as an attentive reader may have noted, the deverbal –Us and deadjectival –(U)Us appear to be formally identical and homonymous. Therefore, a possible objection might be that, in fact, the lack of base frequency effect observed in Experiments 1 and 2 could be explainable by the homonymy of the suffixes alone. Various studies have shown that the processing of truly ambiguous suffixes that denote two (or more) different grammatical functions and/or syntactic categories is affected by homonymy (see Bertram et al. 2000b for Finnish –jA; Bertram et al. 2000c for Dutch –er). However, in contrast to the above suffixes, both the deverbal –Us and deadjectival –(U)Us are syntactically unambiguous (i.e., they are both used to form nouns only). Moreover, their semantic functions are very similar: the former derives words meaning ‘the act of BASE’ and the latter words meaning ‘the state or quality of BASE’, that is, names of actions and qualities, respectively. Although, strictly linguistically, they are two different suffixes (witness the suffix allomorphy of the deverbal –Us vs. deadjectival –(U)Us), to our minds, they can be semantically characterised as having different senses, rather than different meanings and/or syntactic functions. If (un)relatedness of meaning is taken to be a decisive criterion of homonymy, then, strictly speaking, the two suffixes are not homonymous. As there is also recent evidence that even content words with multiple (related) senses have a processing advantage over the truly

9 More precisely, –Us and –(U)Us are homonymous only in nominative singular and in all plural cases except the nominative. Furthermore, words in deadjectival –(U)Us quite rarely occur in any of the plurals, where the overlap of the two suffixes would be greatest. For example, for a frequent noun in –(U)Us, ‘kauneus’ (beauty), that is found 722 times in the Karjalainen database, there is not a single instance of plural inflection.
ambiguous (homonymous) ones (Klepousniotou, 2002), it is more likely that the (partial) homophony/homography of the two semantically and syntactically very similar suffixes would enhance their affixal salience rather than diminish it, as would be the case with truly polyfunctional suffixes.

GENERAL DISCUSSION

The present study adds to the experimental evidence of the processing of Finnish morphology cumulated over the last 10 years. Until now, most studies have indicated that Finnish derived words are generally accessed and recognised by their whole word constitution. The hypothesis of the present paper was that structural invariance of a suffix form would increase the morphological transparency of the complex word and therefore significantly improve affixal salience, the probability of a derivational suffix to serve as a processing unit. This hypothesis was supported by our experiments. Both the denominal collective noun suffix –stO and the deadjectival moderative adjective suffix –hkO showed significant effects of the involvement of base frequency, despite the rather marginal productivity of –stO and the relatively low frequency – in both tokens and types – of –hkO. In contrast, the productive deverbal noun forming suffix –Us and the highly productive deadjectival suffix –(U)Us with rich suffix allomorphy were found to exhibit surface frequency effects only. As other potential confounding factors were ruled out, it seems plausible that having a number of allomorphic variants complicates decomposition for words in these suffixes. The question to be answered, then, is why the decomposition process might run into trouble here.

A decomposition model that makes specific claims about allomorphy is the single direct access model of lexical processing and representation of Marslen-Wilson and associates based on a series of cross- and intramodal priming experiments (Marslen-Wilson, Tyler, Waksler, & Older, 1994; Marslen-Wilson & Zhou, 1999). The basic organisation of the model includes phonologically underspecified lexical entries with no separate level of form-based access representations, and nodes attached locally for each appropriate affix representation. In other words, stem allomorphs are thought to be represented by one phonologically underspecified entry rather than by independent representations for each variant. The model also posits inhibitory connections between each suffix representation but not between different prefixes or between prefixes and suffixes. Although the model accounts for stem allomorphy in English derivation and does not directly deal with suffix allomorphy, it is nevertheless possible to extrapolate from the framework in order to consider the implications of the results obtained here. In other words, taking the notion of under-
 specification to the affix level, one could postulate that all allomorphic variants of an affix are represented by a single phonologically underspecified affix representation.

The predictions of such a model with respect to derived words are rather straightforward. As the model relies on morphological decomposition, one would expect to find base frequency effects, and by assuming stronger links between stem and suffix for high surface frequency words than for low surface frequency words, one would expect to find surface frequency effects for derived words. However, since allomorphic suffixes are comparable to phonologically invariant suffixes, in that in both cases there would be only one suffix representation locally attached to the lexical entry, the same effects should occur for any type of suffix. As we have shown, this is not the case for Finnish derivations, as base frequency effects only occur for phonologically invariant suffixes.

From an English point of view, the idea of phonologically underspecified representations might be a plausible one, since allomorphic variation and the number of unspecified phonemes are limited. In Finnish, however, allomorphy is more abundant and often phonologically irregular. Consider the possible variations of -(U)Us, for example, with phonological/orthographic variants of the suffix as -us, -ut, -ude, -ute, -uks, and -ukse. It is hard to imagine what an underspecified representation for this type of suffix would look like.

Therefore we think that our findings do not undermine the single direct access account put forth by Marslen-Wilson and colleagues, but may, in fact, reflect language-specific differences that call for alternative accounts of lexical organization. 10

However, the question remains why processing a word like ‘heikko+us’ (weakness) with the nominative suffix –Us would be hindered by allomorphic suffixes like –Ude-, –Ut-, –Uks-, and –Ute-. The general answer to this question is that the one-to-one correspondence between form and meaning is compromised. For that reason Schreuder and Baayen (1995) assume that affix allomorphy is among the factors that detracts from the discovery of a salient form-meaning relationship, and, consequently, affects both the age of acquisition and the nature of lexical representation.

10 Also, Finnish stem allomorphy is abundant and often not amenable to straightforward rule-based generalisations. It might therefore come as no surprise, that, in contrast to the wealth of experimental results supporting the phonological underspecification account for the representation of English stem allomorphy, several studies have found evidence for the position that in Finnish stem allomorphs have separate representations in the mental lexicon. This evidence is obtained from a variety of experimental paradigms, such as lexical decision (e.g., Laine, Vainio, & Hyönä, 1999), priming (Järvikivi & Niemi, 2002a), masked priming (Järvikivi & Niemi, 2002b), slips of the tongue (Niemi & Laine, 1992), and aphasia studies (Laine et al, 1995; Laine & Niemi, 1997).
for the affixes with allomorphy. This notion is worked out in many connectionist or network models in which the discovery of morphological structure and its use in processing is an epiphenomenon of form-meaning overlap. For instance, in Bybee's model (1985, 2001) each lexical item is part of a larger associative network, where the lexical items are taken to be interconnected only at the phonological and semantic levels. The connections among the items are seen as a consequence of spreading activation within the (phonological and semantic) network, and generalisations about linguistic properties are schematic and gradient in nature rather than process or rule-based. Morphological generalisations are then seen as an emergent property of the phonological and semantic interconnections between the full forms. As a result, there is no greater qualitative difference between the interrelations of so-called irregular verbs, eat-ate-eaten, and (morphologically) completely regular relations, e.g., work-worked. Thus, analogically, general regularities also emerge from the interconnections of, for instance, send-lend-bend-blend, just as there are regularities between worked-rocked-locked. Applied to the present situation, the model would predict that there are stronger links connecting words with invariant suffixes than words with suffixes that exhibit allomorphy. Although there would be no qualitative difference, the emerging gradient subregularities would make the invariant suffixes relatively more salient units than the suffixes with many allomorphs. However, as we noted earlier, even though there are many allomorphic variants of deadjectival –(U)Us and deverbal –Us, they are realised frequently in the nominative form and, since they are both productive, words in these suffixes would be interconnected extensively in a lexical network as sketched here.

To us it seems therefore that the allomorphic variants introduce some kind of competition element during processing. There are two possibilities. First, there can be some activation feedback mechanism through which allomorphic variants get activated after access of stem and suffix, but before meaning composition and/or other later processes have taken place. The activation of allomorphic variants, even at a later stage of processing, would detract from the computational resources and hinder meaning computation. Second, it is possible that all allomorphic variants get activated early in processing and the incorrect ones have to be suppressed at the very early stages of lexical processing. Also this would require computational resources that complicate processing the complex word via the decomposition route.

The first possibility can be worked out within the framework developed by Schreuder and Baayen (1995), in which three layers corresponding to (A) access representations, (B) concept nodes, and (C) semantic/syntactic nodes are postulated. Activation is taken to flow forward from A via B to
C, but also a feedback activation mechanism is built in. Applying this framework to the situation here, it is possible that the allomorphic variants of –Us and –(U)Us get activated by means of activation feedback, before processing of these words is completed. The assumption would be that all allomorphic variants have separate access representations. This situation is depicted in Figure 1.

As Figure 1 shows, initially there would be only feed forward activation from the access nodes of ‘heikko’ (weak) and –us (–ness) to their concept nodes to the connected semantic and syntactic nodes. However, after the relevant semantic/syntactic nodes have been activated, activation is fed backward via the concept nodes to all the allomorphic access representations. As a result, several access nodes are activated simultaneously. If suppression of these nodes intervenes with operations at later processing stages like meaning computation, the slowdown of the decomposition route and the lack of positive base frequency effects in the experiments

<table>
<thead>
<tr>
<th>Access Representations</th>
<th>Concept Nodes</th>
<th>Syntactic and Semantic Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>heikko</td>
<td></td>
<td>weak</td>
</tr>
<tr>
<td>-us</td>
<td></td>
<td>adjective</td>
</tr>
<tr>
<td>-ude</td>
<td></td>
<td>quality/state (-ness)</td>
</tr>
<tr>
<td>-ut</td>
<td></td>
<td>strength</td>
</tr>
<tr>
<td>-ute</td>
<td></td>
<td>weakness</td>
</tr>
<tr>
<td>-uks</td>
<td></td>
<td>noun</td>
</tr>
<tr>
<td>heikkous</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. A schematic depiction of decomposition of a complex word in –us, ‘heikkous’ (weakness). The suffix –(U)Us has access representations at the form level for all the allomorphic variants, which become activated after activation feedback (depicted with the dotted lines) from the concept level.
with allomorphic suffixes would be explained. As can be derived from Table 7, deadjectival –\textit{Us} is mostly realised with the allomorphic variant –\textit{Ude}-. Taken that the resting activation level is highest for this allomorphic variant, one could speculate that this variant would be least hindered by other allomorphic variants. However, this predicts that words in –\textit{Ude}- would be more likely candidates to undergo smooth decomposition than words in –\textit{Us}. In effect, this prediction was tested in Experiment 5. Despite a somewhat stronger reverse base frequency tendency for the words in –\textit{Ude}- than the words in –\textit{Us}, also here only surface frequency was found to have a reliable effect. This result does cast some doubt on the assumption that the processing of the different allomorphic variants is directly affected by their respective resting activation levels and on the functioning of the activation feedback mechanism in general. It should be noted though that by using this allomorphic variant one has to add an additional suffix (in this case, the genitive case marker), making the complex word tri-morphemic and, as of yet, very little is currently known about the kinds of demands that tri-morphemic words, as opposed to less complex ones, impose on the processing system.

Another explanation for our findings can be that all allomorphic variants get activated early in processing and that the incorrect ones have to be suppressed. An analogue can be made to processing of irregular verbs within the framework of Allen and Badecker (2002). They assume a two-level model of the lexicon (the lexeme level corresponding with Schreuder & Baayen’s level of access representations, the lemma level corresponding to the concept node level and its associated syntactic/semantic level), in which irregular forms have their own lexeme level representation, but a shared lemma level representation. In their study they showed that verb forms like \textit{give} and \textit{gave} do not prime each other and their explanation for this is that there are two contrastive forces involved: at the lexeme (also coined orthographic) level \textit{give} and \textit{gave} suppress one another, whereas at the lemma level, they facilitate each other, leading to no priming overall, that is, they cancel each other out. In addition they showed that when irregular verbs are orthographically dissimilar in the present and past tense (e.g., \textit{teach-taught}), priming does occur. This indicates that the orthographic counterforce was not functional any more, so that the priming effect could show up as a result of the shared lemma presentation. Of course we know from many other studies that orthographically closely related words suppress each other (e.g., Drews & Zwitzerlood, 1995; Grainger, O’Regan, Jacobs, & Seguí, 1989.), but what is important here is that two orthographically and morphologically related words from the same verbal paradigm can compete for activation and that the ‘wrong’ non-intended variant has to be suppressed. A similar mechanism can be proposed for our allomorphs. That is, at access level there may be links
between the different allomorphs and upon attempting to access one of them, the others might get activated as well (and/or they might simply get activated due to orthographic overlap with the target string). In effect, suppressing the ‘wrong’ allomorphic candidates may slow down the decomposition route to an extent, that both base and suffix do not have any functional role in processing the complex words they comprise. In other words, the resources dedicated to processing difficulties at earlier stages would delay or intervene with the computation of the complex word form, or, to be more precise, with the composition of semantic and syntactic information of the stem and the suffix at later stages.

In comparing the two alternatives, something should be said about the gains that the lexical processing system attains when adopting a certain kind of functional arrangement. It is likely that a fast and reliable resolution of ambiguity is amongst the goals the system is striving for. From that perspective it would seem that the system is more inclined to early lateral inhibition among allomorphic candidates than to resolve competition among them after the whole processing cycle is completed. That is, the former arrangement would take the ambiguity ‘out of the air’ at a very early stage instead of letting it linger to the later stages. It should be admitted though that further experimentation is required to get a more precise picture of the mechanisms involved. However, no matter whether allomorphs are activated by a postulated feedback mechanism or whether they immediately compete at access level (or whether there is any other mechanism involved), they clearly disturb the decomposition route. In a dual route framework (Bertram et al., 2000; Schreuder & Baayen, 1995), it would then be argued, that, in effect, the decomposition route ends up being so slow, that a holistic processing route would always be completed first. The abundant surface frequency effects indicate that this seems a reasonable explanation.

CONCLUSIONS

Laudanna and Burani (1995) claimed that salient affixes are more likely to act as processing units in a given language than non-salient affixes. The present paper investigated whether structural invariance of affixes, more precisely, lack of suffix allomorphy, would independently contribute to affixal salience. Experiments 3 and 4 showed base frequency effects for complex words with two suffixes with no allomorphs, whereas in Experiment 1 and 2 only surface frequency effects were found for two productive suffixes with rich suffix allomorphy. It was shown that the results could not be explained by suffix frequency, the relative frequency ratio between the derived word and its base or considerations of affix productivity. Thus, in the light of the present as well as earlier evidence
from Finnish, it is argued that lack of suffix allomorphy (i.e., structural invariance) increases affixal salience importantly.

As Laudanna and Burani (1995) also note, however, accounts of morphological processing should take into consideration language-specific factors modulating the input to the processing system as well as general cognitive constraints of the processing system itself. Therefore it is not obvious that suffix allomorphy should affect the processing of derived words in other languages with different morphological constitution to a similar extent.

Vannest et al. (2002) sought to explain the differences between the processing of derived words in English and Finnish from the distributional properties of the two languages. More specifically, they argued that the fact that Finnish derived words in the majority of cases appear either in inflected and/or multiply derived forms may have tipped the balance of storage and computation in favour of whole word processing as compared with English where derivations appear far more rarely in tri- or quadromorphemic words. As we have seen, this appears to apply even to highly productive bimorphemic derivations with suffix allomorphy, but not to derivations with invariant suffixes. Would suffix allomorphy affect the processing of derived words in languages with less abundant morphology, such as English or Dutch, to a similar extent? Vannest and Boland (1999) found no effect of base frequency for English words in –ation, a suffix with several forms -ation, -ution, -ition, -ion, suggesting that suffix allomorphy may affect morphological processing also in English. However, English derivational morphology is characterized by a division into so-called level I and level II affixes roughly corresponding to a division between non-native and native affixes, with the non-native stock, including e.g., -ion, tending to be generally less productive and less transparent both phonologically and semantically than the native ones. Since Bradley (1979) found only surface frequency to play a role in processing of the English Level I derivations in general, the possible influence of suffix allomorphy cannot be determined. Furthermore, the allomorphy in –ion is also different from, e.g., the Finnish –(U)Us, in that the distribution of the different forms of –ion is morphologically conditioned. In other words a single base takes only one of the possible variants of –ion. The case with the Dutch diminutive suffix appearing as -je, -kje, -tje, -pje, and –etje is similar, in that the choice of the variant is dependent on the phonological constitution of the base word. In contrast, all applicable base lexemes in Finnish paradigmatically exhibit all variants of, e.g., –(U)Us, the choice of the form depending on the case inflection. Furthermore, unlike in English or Dutch, Finnish has a rich system of case inflection, and, in contrast to derivational morphology, the majority of inflectional affixes are transparent (i.e., do not exhibit suffix allomorphy). Thus, their high frequency,
productivity, and lack of allomorphy makes inflectional affixes salient processing units in Finnish and a large number of previous studies has indeed shown evidence for morpheme-based processing of Finnish inflected words (e.g., Bertram et al., 1999; Laine et al., 1995, 1999). In contrast, only very few Finnish derivational affixes seem to be sufficiently salient to be used in the course of processing, and one of the main criteria seems to be that they are formally invariant. For a morphologically less complex language, such as Dutch, the overall morphological constitution does not necessarily create a similar situation. Therefore, for example, the Dutch diminutive suffix –je, may well be frequent and productive enough to actively partake in the processing of words in –je despite the suffix allomorphy.

Specifically what type of architecture is needed to account for the fast accumulating cross-linguistic results remains a task for future research. However, it is clear that – at least in a morphologically rich language – suffix allomorphy modulates affixal salience, and, thus, the processing of derived words. At the same time, an increasingly detailed map of the factors driving decomposition in derivational morphology is being obtained.

Manuscript received November 2002
Revised manuscript received August 2004
PreView proof published online August 2005

REFERENCES


## APPENDIX

### TABLE A
Mean base and surface frequency (per million), base family size (Vf) and length in letters for derivational categories in the base frequency Experiments 1 to 4

<table>
<thead>
<tr>
<th>Affix</th>
<th>Base freq.</th>
<th>Surface freq.</th>
<th>Base Vf</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>dv-Us</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>22.4</td>
<td>2.0</td>
<td>18.6</td>
<td>7.2</td>
</tr>
<tr>
<td>Low</td>
<td>0.5</td>
<td>2.1</td>
<td>12.5</td>
<td>7.2</td>
</tr>
<tr>
<td>da-(U)Us</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>128.1</td>
<td>1.5</td>
<td>23.8</td>
<td>7.1</td>
</tr>
<tr>
<td>Low</td>
<td>8.8</td>
<td>2.1</td>
<td>19.3</td>
<td>6.9</td>
</tr>
<tr>
<td>da-hkO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>74.8</td>
<td>0.1</td>
<td>17.5</td>
<td>8.3</td>
</tr>
<tr>
<td>Low</td>
<td>5.1</td>
<td>0.1</td>
<td>16.3</td>
<td>8.1</td>
</tr>
<tr>
<td>dn-stO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>77.0</td>
<td>0.6</td>
<td>485.5</td>
<td>8.5</td>
</tr>
<tr>
<td>Low</td>
<td>3.3</td>
<td>0.6</td>
<td>61.6</td>
<td>8.5</td>
</tr>
</tbody>
</table>

### TABLE B
Mean base and surface frequency (per million), base family size (Vf) and length in letters for derivational categories in the surface frequency Experiments 1 to 4

<table>
<thead>
<tr>
<th>Affix</th>
<th>Base freq.</th>
<th>Surface freq.</th>
<th>Base Vf</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>dv-Us</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>11.0</td>
<td>9.1</td>
<td>69.4</td>
<td>7.3</td>
</tr>
<tr>
<td>Low</td>
<td>10.8</td>
<td>0.5</td>
<td>20.4</td>
<td>7.3</td>
</tr>
<tr>
<td>da-(U)Us</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>124.5</td>
<td>33.9</td>
<td>114.2</td>
<td>6.9</td>
</tr>
<tr>
<td>Low</td>
<td>119.4</td>
<td>1.6</td>
<td>105.7</td>
<td>7.1</td>
</tr>
<tr>
<td>da-hkO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>89.7</td>
<td>1.0</td>
<td>63.0</td>
<td>8.1</td>
</tr>
<tr>
<td>Low</td>
<td>93.7</td>
<td>0.08</td>
<td>67.0</td>
<td>8.4</td>
</tr>
<tr>
<td>dn-stO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>42.3</td>
<td>2.0</td>
<td>358.9</td>
<td>8.3</td>
</tr>
<tr>
<td>Low</td>
<td>43.2</td>
<td>0.1</td>
<td>308.8</td>
<td>8.6</td>
</tr>
</tbody>
</table>