Advanced students' L1 (Swedish) and L2 (English) mastery of suffixation

MONICA KARLSSON*
Halmstad University

ABSTRACT
Developing the skill to form derivatives is a slow incremental process even for native speakers of English, starting in elementary school and continuing through high school. In fact, it appears to be a universally challenging area of the lexicon. Nevertheless, studies have shown that it is one of the most important skills to possess for a learner aiming to enlarge his/her L2 vocabulary and that it therefore may be worth the while for learners spending time on gaining mastery of derivative forms. In the present investigation, 15 Swedish first-term university students were asked to take two gap-filling context-based tests on suffixation, one in their L1 (Swedish) and one in their L2 (English), both of which were frequency-based (both stem and suffix considered). The students were also asked to evaluate their L1 and L2 knowledge of suffixation. The present study thus addresses the following research questions: Considering 1) the combined frequency of the stem and suffix, 2) the complexity of the suffix (in terms of frequency, predictability, productivity and regularity) and 3) what word classes the suffixes are able to form, what 1) quantitative and 2) qualitative knowledge of suffixation do Swedish university students have in English as their L2 as compared to their L1?

KEYWORDS: second language acquisition, derivation, suffixation, quantitative and qualitative knowledge.

RESUMEN
El desarrollo de la habilidad para formar derivados es un proceso lento y gradual incluso para los hablantes nativos de ingles. Comienza en la escuela primaria y continúa hasta la secundaria. Sin embargo, los estudios han demostrado que es una de las habilidades más importantes que debe poseer un alumno para ampliar su vocabulario de L2 y que, por tanto, invertir tiempo de aprendizaje en el dominio de las formas derivadas puede ser productivo. En la presente investigación, se pidió a 15 estudiantes universitarios suecos de primer semestre que realizarán dos pruebas de rellenar huecos con sufijos basadas en contexto, una en la L1 (sueco) y otra en la L2 (ingles). También se pidió a los estudiantes que evaluaran sus conocimientos sobre sufijación en L1 y L2. Los resultados ofrecen algunas respuestas sobre la relación entre el nivel de conocimientos sobre sufijación en las dos lenguas y factores como la frecuencia de la combinación raíz+sufijo, la complejidad del sufijo y las clases de palabras que los sufijos pueden formar

PALABRAS CLAVE: adquisición de segunda lengua, derivación, sufijación, conocimiento cuantitativo y cualitativo.

*****Address for correspondence**: Monica Karlsson. School of Education, Humanities and Social Sciences, Halmstad University, S-301 18, Halmstad, Sweden; e-mail: Monica.Karlsson@hh.se

1. INTRODUCTION

In the research literature, it seems widely agreed that gaining mastery of derivative forms is a strenuous task. In a study by Schmitt and Zimmerman focusing on advanced learners' production of suffixed words, for example, many incorrect forms produced by the L2 learners appeared. Examples are *releasement* (instead of *release*), *minimizement* (instead of *minimization*), *persistment* (instead of *persistence*) and *survivation* (instead of *survival*) (2002:147). Experiencing difficulty in gaining mastery of derived forms is, however, not only limited to non-native speakers. As a matter of fact, a significant amount of research has also shown that affixation is indeed a challenging task even for native speakers. In Schmitt and Zimmerman's study referred to above, for example, native subjects were included as a point of reference for the non-native speakers investigated. In connection with the native speakers' results, Schmitt and Zimmerman conclude that 'the performance of the native speakers indicates a high but less than complete productive knowledge of derivational morphology' (2002: 160). In fact, gaining derivative knowledge appears to be a slow incremental process for native speakers of English, starting in elementary school (Carlisle, 2000), continuing through high school (Nagy, Diakidoy & Anderson, 1993; Tyler & Nagy, 1989) and, with more infrequent words, it probably involves a lifelong learning situation (Tyler & Nagy, 1990). It also appears to be a universally difficult area of the lexicon to master. In a study on Dutch speakers' knowledge of word formation in their native language, Smedts (1988) found that the seven-year-old informants were only able to provide 14% of the forms sought, that the 13-year-old subjects were only able to give 51% and that, finally, the 17-year-olds included in his study still only mastered 66% of the forms tested.

The question then is: if mastery of word formation is such a slow incremental process even for native speakers, is it worthwhile spending time as a second language learner trying to develop affixation skills? There are three types of studies that point in the direction that it may indeed be the case (Nation, 2001: 264). Firstly, there are quite a few investigations that have focused on exploring the sources of the English vocabulary. These studies show, among other things, that a large number of words can be derived from a limited number of roots. Investigating the 7,476 most frequent words in the Lancaster-Oslo-Bergen corpus (LOB), for example, Bird (1987, 1990) was able to show that 97% of these words could be derived from 2,000 different roots. Focusing on the affixes forming these words thus seems to be time well spent. Secondly, there are studies that have investigated the proportion of affixed words in different corpora. In a study by Cunningham (1998), it was shown that affixed forms (derivational and inflectional) outnumbered stems four to one. In another study by Nagy & Anderson (1984), the ultimate aim was to find out how many word families can be found in all printed school English. They were able to show that as much as 12.8% of the different word family types seen in their study contained derivational affixes (21.9% contained inflected forms), again indicating that derivatives make up a comparatively large part of the

English lexicon. Finally, there are a number of investigations that have focused on the frequency of specific affixes. White, Power and White (1989), for example, were able to show that 60% of all the words containing the prefixes *un-*, *re-*, *in-* and *dis-* could be understood once the learners knew the most common meaning of the base word. If the derived words were contextualized and knowledge of some of the less common meanings of the prefixes was gained, as much as 80% of the derivatives could be understood. Other investigations focusing on the frequency of certain affixes (e.g. Becker, Dixon & Andersson-Inman, 1980; Harwood & Wright, 1956) generally show that a small number of affixes are indeed very frequent and make up a very large percentage of all affix use.

There thus seems to be a great deal of evidence suggesting that it may be useful to develop affixation skills when learning the English language (and, as the result section will show, it seems equally important to acquire affixation skills in Swedish). As a matter of fact, studies have shown that this is one of the three most important skills to possess for a learner aiming to enlarge his/her L2 vocabulary, explicit teaching and learning vocabulary from context being the other two (Nation, 2001: 263). This is of course especially the case with advanced learners who have already developed a comparatively rich vocabulary (Schmitt & McCarthy, 1997).

Also, since the present study deals with advanced learners, the present author decided to focus on derivational suffixation which is normally considered to be the most difficult aspect of word formation (Nagy, Diakidoy & Anderson, 1993: 156). This is, for example, mirrored in the fact that these affixes are generally acquired later than other types of affixes by L1 as well as L2 learners (Berko, 1958; Derwing & Baker, 1979).

2. THEORETICAL ANCHORAGE AND PREVIOUS RESEARCH

2.1. Do learners decompose morphologically complex words or not?

For quite some time researchers have discussed whether the term *word family* may only be used as a linguistic term for a set of words that all have the same base (e.g. *persist*, *persists*, *persisted*, etc.) or if the concept paints a picture of how words really are stored in a learner's mental lexicon. According to Nagy *et al.* (1989) there are several possible approaches to how morphologically complex words may be stored. One possibility is that all of the members of a specific word family can be found under the same lexical entry. It could also be argued, as a possibility at the other end of the continuum, that all the members of a word family are listed under totally separate entries. Between these two extremes, there are yet other possibilities, one being that all the members of a word family have separate entries but that these entries are linked in some way so that if one of the members is encountered, the links between the family members will automatically heighten the learner's awareness of the other members of that same word family. It is also suggested that it is likely that only regular inflections and those

affixes that are semantically transparent can be found under the same entry as the stem, whereas those inflections that are irregular and affixes whose meanings are semantically blurred have separate entries (e.g. Caramazza *et al.*, 1988). Other factors also need to be considered, such as the frequencies of the stem, inflection and derivational affix and whether the spoken and written forms of the stems and affixes are easily recognizable. The age at which, for example, a suffixed word is learned may also decide whether a word is decomposed or approached as an unseparated whole (Gardner 2007:258). It seems that the earlier a word is learned, the more likely it is that it is not decomposed into parts later in life.

In order to test if words really are stored in word families, Nagy *et al.* (1989) set up an experiment with the focus of finding out whether the frequency of the stem alone or if the inflected and derivative forms of that stem also affect the time it takes a person to recognize the stem. For example, is it only the frequency of the word *quiet* as the stem of the word family in question that affects the time it takes for a person to recognize that stem or do the frequencies of the word family members *quietness* and *quietly* also play important roles? If it could be shown that the frequencies of inflected and derived forms play an equally great part in recognising a stem as the frequency of the stem itself, it would indicate that the members of a specific word family can be found under the same entry. If, on the other hand, the frequencies of inflected and derived forms could be shown to have no part in the recognition of the stem, the members of that word family would probably be listed under different entries than the stem.

Nagy *et al.* (1989) conclude that especially the accuracy, but also the speed with which the subjects performed the task, rules out the possibility that every member of a word family has its own separate entry. The opposite position, i.e. that all the members of a word family share the same entry, also appears to be false in that the effect of the frequency of the morphologically related words, though significant, was not as strong as the effect of the frequency of the stems themselves in their experiment. Nagy *et al.* (1989) thus claim that there appears to be partial activation in that some members affect the accuracy and speed of recognition more than others. Regularity appears to be a factor, among others, to consider here. Bradley (1979), for example, saw in his investigation that the frequency of the stem affected the reaction time for suffixed words in a positive way when the spelling and/or pronunciation of the stem involved few changes (as is the case when suffixes like -*er*, -*ness* and -*ment* are added). When the spelling and/or pronunciation involved major changes, on the other hand, the frequency of the stem had no effect (as is, for instance, the case when -*ion* is added).

Recent research also seems to support Nagy *et al.*'s conclusion. In Bertram, Baayen & Schreuder (2000) and de Jong, Schreuder & Baayen (2000), it was shown that the number of derivatives in a word family may affect how those members are stored. The more derivatives, the more the links between the various family members are heightened when a member is activated. This appears to happen both when the activated word is a noun (Schreuder &

Baayen, 1997) and a verb (de Jong, Schreuder & Baayen, 2000), in students' first as well as second language (Mulder, Schreuder & Dijkstra, 2013).

Another study relevant to the present section is Schmitt and Zimmerman (2002). According to this article, previous research shows that there is psychological justification for the term *word family*. It has also been shown, Schmitt & Zimmerman write, that learners are able to recognize unknown members of a word family from a member they already know, i.e. there is evidence that the organization of words into word families helps learners' receptive knowledge. What has not been investigated thoroughly yet, according to Schmitt and Zimmerman, is learners' productive knowledge of other members of a word family once one of the members becomes known. For example, if a learner demonstrates knowledge of the verb *stimulate*, can he/she then also automatically produce the inflected forms *stimulates* and *stimulated* and the derivative forms *stimulation* (noun) and *stimulating* (adjective)?

The investigation performed by Schmitt and Zimmerman addresses three research questions. Firstly, they wanted to find out how many of the four word classes formed with the help of affixes (noun, verb, adjective and adverb) that the learners knew productively. They also wanted to find out which of the word classes the subjects tended to know the best. Thirdly, they wanted to investigate the relationship between productive derivational knowledge of a word and more general knowledge of that same word. Here the subjects were asked to give information concerning the degree to which they knew each word tested, much in the same way that the students in the present study were asked to provide information about their knowledge of the derivatives (and stems) tested (see Section 3).

The results of the study show that being able to produce other members of a word family once one of the members has been learned is no easy task. The students in Schmitt & Zimmerman's study managed to give 58.8% correct answers, which means that on average the subjects were able to produce two of the four derivatives for a given word family (with a difference between advanced students who produced 3-4 forms and the less advanced students who produced 2-3 of the 4 forms). Furthermore, all four forms were produced correctly in only a few cases. On the other hand, there were also only a few occurrences in which the subjects were not able to produce any correct forms. Finally, of the four word classes, verb derivatives were best known (67%), followed by noun derivatives (63%), adjective derivatives (54%) and, lastly, adverbs (52%), proving to be the most difficult word class to deal with. One reason why verbs appear to be the easiest word class may be that the verb is often the stem of a word family and is thus the form encountered more frequently than the others.

The results also showed that the subjective information given by the students concerning the degree to which they thought they knew the word in question gave a good indication of whether they would get the word right or not. This was especially true for verbs and nouns, separating the two more semantically transparent word classes from the two that are less so (adjective and adverb). Also, in addition to producing incorrect solutions for which

the subjects stated that they knew the derivative, it is interesting to note that learners also produced quite a few derivatives which they claimed were fully unknown to them.

Finally, even if all of the native speakers claimed that they knew all of the derivatives tested they did not show complete mastery when asked to produce them. Nevertheless, they were able to produce the correct forms in 91.8% cases, thus outperforming both non-native university and pre-university students, where the former in turn outperformed the latter.

There are also a number of other studies that point in the same direction as does Schmitt and Zimmerman's investigation. For example, in Schmitt and Meara's longitudinal study from 1997 the non-native university students tested were only able to produce 15% of all the derivatives asked for. In yet another longitudinal study, Schmitt (1998) included three postgraduate university students. Even these very advanced students undertaking PhD studies had clear gaps in their production of derivative forms and when tested a second time they showed no great improvement. Finally, in Schmitt (1999) the subjects, also university students, could only produce 12 of the 180 derivatives (nouns, verbs, adjectives and adverbs) asked for. However, these three studies differ in one respect from the Schmitt & Zimmerman study in that the subjects were asked to provide the derivative forms out of context. This probably affected the results in a negative way.

The studies discussed above all help to shed light on the real issue at hand: when a learner wants access to a specific derivative does he/she then take the decomposition route or the direct route, i.e. does he/she analyze the derivative into parts or does he/she analyze it as an unseparated whole? In fact, more and more research seems to indicate that learners start their search by doing both and that the route finally made use of is the one that takes the shortest time to gain the sought-for answer (Haspelmath & Sims, 2010). There are several factors that decide which route will be the fastest one. One such factor is certainly the frequency of the stem in relation to the frequency of the derivative. Learners appear to have an innate knowledge of frequencies of words and can thus subconsciously decide if it is the stem or the derived form that carries more 'memory strength.' This means, for instance, that if the derivative is more frequent than the stem, it is very likely that the learner in such a case would end up using the direct route to gain lexical access to the word in question. An example of a derivative which probably works in this way is the adjective *insane*. This derivative is more frequent than the stem *sane* (i.e. *insane* has more memory strength than *sane*) and will thus not be decomposed when sought for in the mental lexicon. If, on the other hand, the stem carries more memory strength than the derivative, i.e. the stem is more frequent than the derivative, it is very likely that the learner encountering such an item will take the decomposition route. An example of a word that probably works in this way is *modernity* in that the stem *modern* is more frequent than its derivative form (Haspelmath & Sims, 2010: 123). Although the frequency of a word appears to be the main contributor to whether a word has great memory strength or not (as discussed above), there are other factors too that may play a role. For example, a word may carry great memory strength because it sounds or looks

funny/odd etc. or because it has some sort of other personal appeal to a specific learner. Thus low-frequency words may sometimes also have great memory strength. One such example is the word *nincompoop*, the meaning of which the present author never forgot after the first encounter.

As discussed in connection with the Nagy *et al.* (1989) study above, affixes that induce allomorphy in the base also affect which route a learner finally takes when trying to gain lexical access to a derived form. The English suffixes *-ity* and *-ship* may serve as examples of two opposites. Whereas *-ity* causes allomorphy to occur in the base (as in *electric – electricity* and *divine – divinity* which both involve changes in pronunciation), there is no such change in the base when *-ship* is added (as in *ambassador – ambassadorship*). This may mean that when these derivatives are sought for in the mental lexicon, *electricity* and *divinity* are decomposed, whereas *ambassadorship* may be accessed through the direct route. Put differently, since *ambassadorship* is more easily segmentable than *electricity* and *divinity*, a learner will most likely access *ambassadorship* via the direct route whereas *electricity* and *divinity* will be accessed via the decomposition route (Haspelmath & Sims 2010:72-74).

Furthermore, an extension of the supposition that it is the relative frequencies of the stem and the derivative that decide whether a word is decomposed or not is that the affix itself is also affected by which of the two processes is eventually chosen. If an affix tends to occur in words that are decomposed, the lexical entry of such an affix will be frequently accessed and such an affix will therefore carry greater memory strength (referred to as the parsing ratio of an affix). According to Haspelmath and Sims (2010:124), this means that such an affix is likely to have high productivity, i.e. it is more readily available for speakers when coining new words. From a second language learner perspective, it would be interesting to see if affixes with great memory strength are also those affixes with which learners have the least problems and/or make use of the most if they are uncertain of what affix to produce. Table 1 shows the parsing ratio for a number of English affixes. The higher the parsing ratio, the more the affix is activated in the lexicon, i.e. such affixes are very often seen in words that are decomposed. Thus, whereas *-ence* is a suffix that appears to occur in words that are primarily stored as unseparated wholes, *-less* is a suffix that mostly occurs in words that are decomposed and, in theory, this means that the suffix *-less* should be more easily learned than the suffix *-ence*.

Suffix	Parsing ratio
-ence	0.1
-ity	0.17
-ate	0.31
-dom	0.5
-ness	0.51
-ish	0.58
-like	0.68
-proof	0.8
-less	0.86

Table 1. The parsing ratios for a number of suffixes
(Haspelmath & Sims, 2010: 124, in turn from Hay & Baayen, 2002: 233-235)

2.2. A pedagogical approach to affixes

In order to provide teachers with a more systematic way of approaching affixes, Bauer and Nation attempt in their article 'Word Families' (1993) to divide affixes into increasing levels of difficulty. There are four criteria upon which Bauer & Nation base these levels: 1) frequency, i.e. the number of words in which the affix appears, 2) predictability, dealing with the degree of transparency of the meaning of the affix, 3) productivity, i.e. the extent to which the affix in question is able to form new words, and 4) regularity, considering both the base to which the affix is added and the affix itself (Bauer & Nation, 1993: 255). In all cases priority is given to enhance students' understanding of written texts, thus the four criteria are mainly based on the word-building process of the written form. Additionally, Bauer and Nation are careful to explain that their difficulty levels are not discrete or absolute, but rather form a continuum. (No corresponding continuum could be obtained for Swedish suffixes.)

The continuum based on these four criteria comprises seven levels in which the strictness of the inclusion of the criteria decreases incrementally as one approaches the higher levels. Levels 3-7 are of special interest to the present study since they deal with derivational suffixes. Examples of suffixes belonging to Level 3 (the most frequent and regular affixes) are: *-less* (*clueless*), *-ly* (*charmingly*) and *-ness* (*commonness*); Level 4 (frequent and orthographically regular affixes): *-ation* (*realization*), *-ity* (*reality*) and *-ment* (*accomplishment*); Level 5 (regular, but infrequent affixes): *-al* (*arrival*), *-ance* (*clearance*) and *-ence* (*emergence*); Level (frequent but irregular affixes): *-ive* (*supportive*) and *-th* (*warmth*); Level 7 (classical roots and affixes): *-ate* (*captivate*) and *-ure* (*departure*). (See Bauer and Nation (1993) for more examples of suffixes at the different levels.) It should also be noted that one and the same suffix, but with different meanings, may occur at more than one level.

One drawback with this framework is that it does not address the role played by the stem (Gardner, 2007). As mentioned previously, the frequency of the stem appears to be an important factor to consider when discussing the ease or difficulty with which learners are able to produce derivative forms. Other investigations point in the same direction. For

example, there are studies that show that L1 students tend to focus entirely on the meaning of the stem when asked to define a derivative (e.g. Wysocki & Jenkins, 1987).

3. THE STUDY: METHOD

3.1. The informants and research questions

In the present investigation, 15 Swedish first-term university students, having studied English for about ten years, were asked to take two gap-filling context-based tests, one in their L1 (Swedish) and one in their L2 (English), both of which were frequency-based (both stem and suffix considered). There were 12 female and 3 male students, 10 of whom were in their late teens or early 20s, 4 in their late 20s and 1 49-year old (range: 19-49; mean: 23.47; SD: 7.68). The informants were also asked to evaluate their L1 and L2 knowledge of suffixation. The present study thus addresses the following research questions:

> Considering 1) the combined frequency of the stem and suffix, 2) the complexity of the suffix (in terms of frequency, predictability, productivity and regularity) and 3) what word classes the suffixes are able to form, what 1) quantitative and 2) qualitative knowledge of suffixation do Swedish university students have in English as their L2 as compared to their L1?

Also, one native speaker, a 33-year old male studying within the Swedish educational system to become a high school teacher of English, was used as a point of reference for the L2 test.

3.2. The tests

Since English and Swedish have a similar capacity for forming derivatives (Andersson, 1987: 26), two identically structured tests could be put together, both of which are context-based gap-filling tasks. 50 items, all of which were picked randomly from various books/booklets containing extensive lists of stems and their derivative forms, were tested in each language (all test items can be seen in Tables 3 and 4). Furthermore, the L2 test preceded the L1 test, which was taken directly after the English test with only a short hiatus in between. Also, the students were able to sit with each test as long as they liked, i.e. no time constraints what so ever were put on the test in either language.

The context was mainly taken from the same books/booklets as the test items. In most cases the students were obliged to add a suffix, such as in the two examples below:

> (1) The bus **arrived** late. The estimated time of _____ was 10.00 hrs.

(2) Don't **encroach** on my rights!   Your _____ is unforgivable.

In a few cases, though, the students were asked to give the stem, i.e. they had to remove the suffix from the word given, as in:

(3) I like being **comfortable**.   I love _____ .

Furthermore, on both tests the students were also asked to evaluate the degree to which they believed they knew the words, both the items given and the items the students were supposed to produce. For the given word, they were to decide whether they did not know the word at all, if they thought they knew the word or if they were sure that they knew the word. For the word they were expected to fill in they were asked to make a decision along the same lines. This is exemplified below.

(4) The drawing has many **colours** in it.   It is very _____.

☐ I don't know the word.   ☐ I don't know the word to fill in.
☐ I think I know the word.   ☐ I think I know the word to fill in.
☐ I am sure I know the word.   ☐ I am sure I know the word to fill in.

For each correct answer, the students were awarded one point, thus a total of 50 points on each test could be achieved. As in Schmitt and Zimmerman (2002: 157), minor misspellings were ignored as long as the intended stem/derivative could be discerned.

Also, the test items were presented to the students in order of frequency based on the BNC and Språkbanken (a Swedish corpus) respectively. Moreover, since, as discussed in Section 2, there appears to exist ample evidence that stems and their derivatives are often stored together in a learner's mental lexicon and that their respective frequencies interact when it comes to whether a derivative is decomposed or not, it is the sum of the frequencies of the members involved in the word pairs that has determined the order in which they were presented to the students. The total frequency of the test items was also considered. Here the English items were seen to be slightly more frequent than those included on the Swedish test, which means that in terms of frequency the English test was slightly easier than the Swedish one.

Lastly, in order to make the following discussion concerning the students' results as easy to read as possible, the present author will sometimes refer to the words formed as derivatives/derivative forms although some of these are in fact stems that have been formed by truncating the related derivatives.

4. RESULTS AND DISCUSSION

4.1. The students' mean scores

Table 2 presents the students' mean scores on the two tests. It also includes the native speaker's result.

Students	Mean (50 items)	Standard deviation	Standardised scores
Native speaker of English	30.0	-	-
Swedish students – English test	27.6	6.29	highest: 1.33 lowest: -2.00
Swedish students – Swedish test	42.9	4.86	highest: 1.45 lowest: -2.45

Table 2. The results on the suffixation test

Unsurprisingly, the Swedish students did considerably better on the Swedish test (with the mean score of 42.9) than on the English test (mean score: 27.6). Nor is the native speaker's result very surprising. Even though it is comparatively low (30.0), he still, as expected, outperformed the Swedish students. This native speaker achieved poor results on all of the linguistically oriented courses he took for the present author. Through personal communication with the lecturer teaching literature courses, the present author also learned that he was a low-achiever in those classes too. It is therefore likely other natives would have produced better results.

4.2. Is there a frequency effect on students' suffixation skills?

Figures 1 and 2 present the number of correct answers for each of the 50 test items in the English and Swedish test respectively. The reader is reminded that the items were presented to the students in order of word-pair frequency so that whereas derivative number one is part of the most frequent word pair, derivative number fifty belongs to the most infrequent word pair. Furthermore, the derivatives whose scores are indicated in bold in Figure 1 are part of West's *A general service list of English words* (1953), listing 2,000 head words of high frequency. No such corresponding word list could be found for Swedish words.
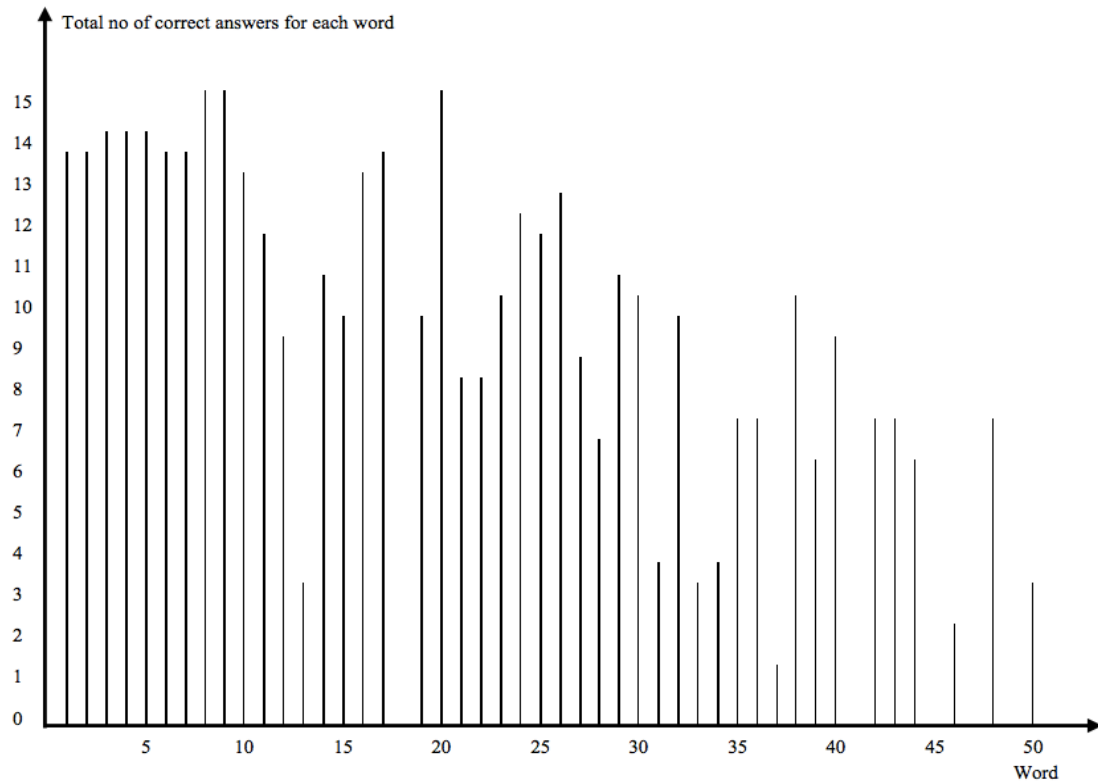
Figure 1. The total number of correct answers for each word on the English test
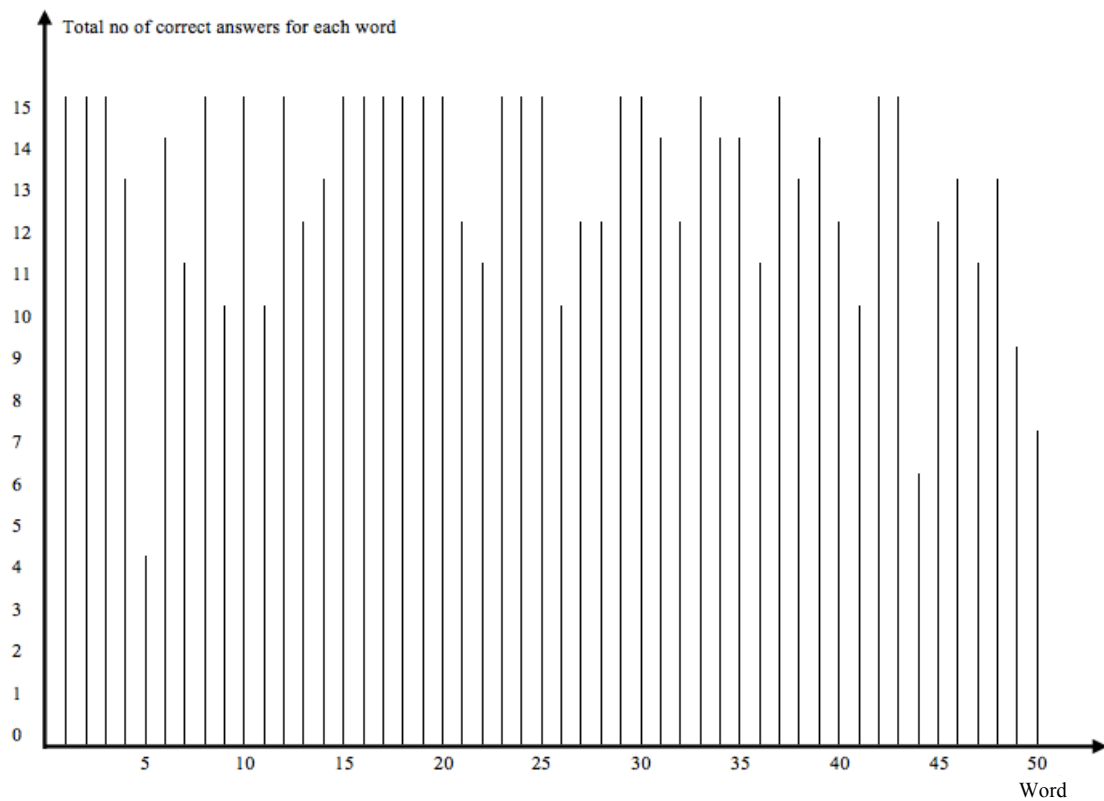

Figure 2. The total number of correct answers for each word on the Swedish test

In Figure 1 it can be seen that word-pair frequency does have an effect on the students' results on the English test in that, with a few exceptions, there is a clear decrease in the number of correct answers from the most frequent word pair to the least frequent one. It is also interesting to note that most of the words that are a part of the first two thirds of these fifty words are also found on West's list, thus solidifying the word-pair frequency effect further. This result can be compared to the native speaker's achievement which does not appear to be affected by frequency at all. For the first 25 test items, he was able to produce 16 correct derivative forms and for test items 26-50, he offered 14 correct derivations. On the corresponding Swedish test part, just as with the native speaker's result on the English test, no clear decrease in relation to word-pair frequency can be seen (Figure 2) and this is despite the fact that the Swedish test items are slightly more difficult than the ones on the English test, at least if difficulty is measured in terms of frequency. It is not until the very last few word pairs that a slight decrease might be detected.

4.3. The two suffixation tests in more detail

Tables 3 and 4 present the students' results for each word on the English and Swedish test respectively (translations of the Swedish items are offered in the Appendix). In both tables, the words formed by the students are presented in order of accuracy so that the word that received the highest score is placed first and the word with the fewest correct answers placed last. When more than one derivative received the same score, the word for which the students showed the greatest accuracy in their self-evaluation is placed first. This is explained further below.

In Table 3 (English test), words included on West's general service list (1953) are, as in Figure 1, indicated in bold. In addition, in both tables the students' evaluation of their knowledge of the given word as well as the word to be formed is included. For example, for word pair number 3 (*apply – application*), for which the number indicates where on the test the word pair occurs and consequently also its frequency in relation to the other word pairs, the table shows that *apply* (which is given in the past tense on the test, indicated in parenthesis after the word's base form) is the word given and *application* the derivative the students were meant to form. 14 students gave the correct derivative (the number of correct answers is indicated immediately after the word to be formed), thus only 1 failed to do so. 12 of the students were *sure* that they knew the meaning of the given word (column KN(A)) and 9 of the students were *sure* that they had offered the right derivative (column KN(B)). 2 of the 15 students were less sure of the meaning of the given word in that they indicated that they only *thought they knew* the meaning of *applied* (column TK(A)). Both students were here correct in their assumption. 5 students indicated that they *thought they knew* what derivative to fill in (column (TK(B)), 1 of whom was wrong in this assumption (indicated in parenthesis). None of the students stated that they did not know the meaning of the word given (column NK(A)) or that they did not know what derivative to write (column (NK(B)). For this specific word

pair, 14 of the 15 students offered the requested information (the sums of the figures in columns KN(A) and TK(A) on the one hand and KN(B) and TK(B) on the other hand both yield 14 answers respectively). This means that, in the example discussed, the majority of the students who answered the requested information knew both the meaning of the word given (KN(A)) and what word to form (KN(B)). In contrast, for word pair 42 (*abbreviate – abbreviation*) the majority of the students (10) did not know the meaning of the word given (KN(A)), nor did the majority know what word to form (KN(B)).

Also, in a few cases more than one derivative form was considered to be correct. For instance, both *intrepidity* and *intrepidness* were accepted as possible noun derivatives of the adjective stem *intrepid*. Similarly, both *ruminative* and *ruminant* were considered to be correct adjective derivatives of the verb *ruminate* and thus both awarded a point. The students' scores for these alternatives are indicated in parenthesis immediately after each derivative form. The present author based her acceptance/rejection of such alternatives on the presence/absence of them in the BNC. Alternative forms also exist in the Swedish material for which the corresponding corpus was used to decide whether they would be included or not.

NO	WORD GIVEN A	WORD FORMED B	KN(A)	KN(B)	TK(A)	TK(B)	NK(A)	NK(B)
9	advise(d)	advice 15	13	11	1	3		
8	colour(s)	colourful 15	13	9		4		
20	anxious	-ness (8), anxiety (7) 15	11	9	3	4		1 (1)
4	bear (born)	birth 14	13	12	1	2 (1)		
5	enter	entrance (11), entry (3) 14	13	10 (1)	1	4		
3	apply (applied)	application 14	12	9	2	5 (1)		
10	achieve	achievement 13	13	12	1	2 (2)		
17	arrive(d)	arrival 13	13	8	1	4 (1)		2 (1)
6	deep	depth 13	12	6	1	7 (2)		
7	easily	ease (9), easiness (4) 13	12	5	2	7		2
2	early	earliness 13	12	2	2	12 (2)		
16	clean(ed)	cleaner('s) 13	11	2	3	9 (1)		3 (2)
1	able	ability 13	10	7	3	5		1
24	clever	cleverness 12	12	3 (1)	2	9 (1)		2 (1)
26	cautiously	-ness (2), caution (10) 12	11	8 (1)	2	4	1	2 (1)
11	equal	equality 11	13	10 (1)	1	4 (2)		
25	cruel	cruelty 11	12	9 (1)	2	4 (1)		1
29	envy (envied)	envious 10	12	7	1	6 (4)	1	1
14	absent	absence 10	11	5	3	8 (3)		1
23	courage	courageous 10	11	6 (1)	3	6 (2)		2 (1)
30	acquaint(ed)	acquaintance 10	9	7 (1)	4	5 (1)	1	2
38	inordinate	inordinately 10	2	1	5	7 (3)	7	6 (5)
32	annoy(ed)	annoyance 9	12	4 (1)	2	8 (3)		2
15	comfortable	comfort 9	12	5 (2)	2	8 (4)		1 (1)
12	advantage	advantageous 9	11	3	2	7 (2)		3
19	broad	-ness (9), breadth (0) 9	10	2 (1)	3	7 (3)	1	5 (3)
40	intrepid	-ness (2), intrepidity (7) 9			4	6 (1)	10	8 (4)
27	admire(d)	admiration 8	13	7 (1)	1	6 (4)		1
22	complain(ing)	complaint(s) 8	12	8 (4)	2	6 (3)		
21	bury (buried)	burial 8	11	5 (1)	2	6 (2)	1	3
36	contagious	-ness (7), contagion (0) 7	10	3 (2)	2	5 (3)	1	5 (4)
42	abbreviate	abbreviation 7	4	4		2 (2)	10	8 (2)
35	encroach	encroachment 7	3	2 (1)	2	4 (2)	9	8 (4)
43	surly	surliness 7	1	1	4	3 (1)	9	10 (4)
48	ruminate	ruminative (5), ruminant (2) 7			3	5 (2)	11	9 (4)
28	(did)n't approve	disapproval 6	13	9 (4)	1	5 (4)		
44	voracious	-ness (4), voracity (2) 6	1		4	7 (2)	9	7 (1)
39	exacerbate(d)	exacerbation 6	1		1	5 (3)	12	9 (4)
13	eat	edible 3	13	6 (3)	1	6 (6)		2
34	coward	cowardice 3	13	4 (2)	1	8 (7)		2

NO.	WORD GIVEN A	WORD FORMED B	KN(A)	KN(B)	TK(A)	TK(B)	NK(A)	NK(B)
31	**conquer(ed)**	**conquest 3**	12	6 (3)	2	6 (6)		2
33	complacent	complacence[1] (0) 3	2	1	2	4 (2)	9	8
50	asinine	asininity 3			2	5 (3)	12	9 (1)
46	guileful	guile 2		1 (1)	3	4 (3)	11	9 (1)
37	tempest	tempestuous 1			5	4 (4)	9	10 (1)
18	capacity	capacious 0	11	4 (4)	3	6 (6)		4
47	chastise(d)	chastisement 0			4	6 (6)	10	8
49	putrefy	putrefaction 0			4	6 (6)	10	8
41	arcane	arcanum 0			3	5 (5)	9	7
45	parsimonious	parsimony 0			2	5 (5)	12	9

Table 3. The students' results on and evaluation of the English suffixation test, the table listing the test items in order of the number of correct answers starting with the item that received the highest score. (KN=the word is <u>kn</u>own, TK=the word is <u>t</u>hought to be <u>k</u>nown, NK=the word is <u>n</u>ot <u>k</u>nown)

NO.	WORD GIVEN A	WORD FORMED B	KN(A)	KN(B)	TK(A)	TK(B)	NK(A)	NK(B)
4	allvarlig	allvar 15	14	14				
19	cykel	cyklist 15	14	14				
18	glömt (glömma)	glömska[2] (12) 15	14	13		1		
24	mätt	mätthet (2), mättnad (13) 15	14	13		1		
30	brutal	brutalitet 15	14	13		1		
43	tolerera(r)	tolerant 15	14	13		1		
2	vecka	veckovis 15	14	12		2		
9	organisation(en)	organsiera 15	14	12		2		
17	arrangera(t)	arrangemang 15	14	12		2		
16	journalist	-ik (11), journalism(4) 15	14	11		3		
21	teori(n)	teoretisk 15	14	10		4		
25	skam	skamfull (1), skamsen (14) 15	14	10		4		
29	blyg	-het (14), -sel (1) 15	14	10		4		
33	skryt(a) (-er)	skrytsam 15	14	9		5		
1	förslag	förslagsvis 15	14	8		6		
37	vilsen	vilsenhet 15	14	8		6		
3	centralt	centrum 15	13	14	1			
13	acceptabel(t)	acceptera 15	13	13				
20	fattig(a)	fattigdom 15	13	13				
11	erbjud(a) (-er)	erbjudande 15	13	12		1		
42	enfaldig	enfald (2), enfaldighet (13) 15	13	9	1	5		
7	uppenbar(t)	uppenbarligen 14	14	13 (1)		1		
39	förälska(r) sig	förälskelse 14	14	12		2 (1)		
34	idiot	idioti 14	14	10		3		1
31	komfort(en)	komfortabel 14	14	10		4 (1)		
35	staty	statyett 14	14	7		6		1
15	dansa	dansare (10), dansös (3) 13	14	13 (1)		1 (1)		
48	konfirmera sig	konfirmand 13	14	12 (2)		2		
5	köra (kört)	körbar (4), körd (9) 13	14	9 (2)		5		
46	verifikation	verifiera 13	12	9	2	5 (2)		
38	teologi	teolog 13	11	8 (1)	1	5 (1)	1	
32	tåla	tålig(a) 12	14	10 (1)		4 (1)		
14	medveten	medvetenhet 12	14	11 (2)		3 (1)		
22	tvivel	tvivelaktig 12	14	8 (2)		6 (1)		
45	ståupp	ståuppare 12	14	6		8 (3)		
27	stipendium	stipendiat 12	14	6		5 (2)		3 (2)
40	fanatism	fanatiker 12	13	11	1	3 (2)		
28	feber(topp)	feberfri 12	13	10 (2)		3		
36	hysterisk	hysteri 11	14	11 (3)		3		
23	sanning(en)	sanningsenlig 11	14	5 (1)		9 (2)		
8	aktuell(t)	aktualisera 11	13	5 (1)		7 (1)		1
26	lejon	lejoninna 10	14	11 (4)		3 (1)		
47	instruera	-(a)nde (2), instruktiv (8) 10	14	5		9 (2)		
12	kosta(t)	kostsam 10	13	10 (2)		3 (2)		

41	kvantitet(en)	kvantitativ 10	12	8 (1)	1	2 (1)	1	4 (2)
10	(hög)säsong	säsongbetonat[3] (4) 10	11	2	1	6 (1)		4 (3)
49	fabrikation	fabricera 9	9	6 (2)	4	8 (3)	1	
50	gondol	gondoljär 7	12	2	2	8 (3)		4
44	ensligt	ensling 6	13	7 (3)	1	5 (4)		2
6	avtal(et)	avtalsenlig(t) 4	13	5 (4)		8 (5)		

Table 4. The students' results on and evaluation of the Swedish suffixation test, the table listing the test items in order of the correct number of answers starting with the item that received the highest score. (KN=the word is <u>kn</u>own, TK=the word is <u>t</u>hought to be <u>k</u>nown, NK=the word is <u>n</u>ot <u>k</u>nown)

4.3.1. The results on the English test

The students' results on the English test can be divided into three main sections, each section having different characteristics. The derivatives that received scores from 15 through 10 (*inordinate – inordinately* excluded), i.e. from *advise(d) – advice* (word pair 9) through *acquaint(ed) – acquaintance* (word pair 30), form one such group. Firstly, these twenty-one word pairs can all be found among the thirty most frequent word pairs. As shown in Figure 1, the majority of these derivatives are all found on West's general service list. As can be seen, it is however not only most of the *derivatives* that are found on West's list, but also the words on which these derivatives are based, i.e. the majority of the items in all these word pairs are frequent words that these Swedish learners probably have encountered quite a few times during the ten years or so that they have been studying English. In Sections 1 and 2, it was discussed whether derivatives are decomposed or regarded as unseparated wholes when encountered by a learner and that age may be one of many factors determining whether decomposition occurs or not. Marslen-Wilson et al. (1984), for instance, claim that the younger a learner is when encountering a derived form, the less likely it is that the learner decomposes it when encountering it later in life. Due to their frequency, these word pairs may thus be tackled as unseparated wholes. The parsing ratios for two of the suffixes involved here also appear to point in this direction. The fact that the suffix *-ness* (as, for instance, in *cleverness* (word pair 24)) has a much higher parsing ratio than *-ity* (as in *equality* (word pair 11) (see Table 2)), does not appear to affect the students' performance with these derivatives to a great extent, even though this difference in parsing ratios means that the students should have encountered the suffix *-ness* more often than the suffix *-ity*, since the former is decomposed more often than the latter. Within this group, however, it should be noted that the derivatives ending in *-ness* generally received higher scores than the ones ending in the suffix *-ity*.

Furthermore, the students' ability to produce correct derivatives for these word pairs seems to be independent of the complexity of the suffix in question (in terms of Bauer and Nation's approach presented earlier), a fact that also indicates that the students approach these derivative forms as unseparated wholes. For example, the students are not only able to form the derivatives *anxiousness*, *ability/equality* and *courageous* with the suffixes *-ness* (Affix Level 3 which includes the most frequent and regular affixes), *-ity* and *-ous* (Affix Level 4 which also includes comparatively frequent affixes that are still quite regular), but also the

derivative forms *entrance, arrival, application* and *birth/depth* with the suffixes *-ance*, *-al*, *-ion* and *-th*. These latter suffixes, in contrast to *-ness* and *-ful*, belong to the more complex Affix Levels 5 (*-ance* and *-al*) and 6 (*-ion* and *-th*), where Level 5 includes regular but infrequent affixes and Level 6 deals with comparatively frequent affixes but whose spelling entails difficulties. If the students approach these derivatives as unseparated units, which seems to be the case, it may mean that they have not understood what these suffixes do (i.e. their meaning) and consequently not what word classes they are able to form, i.e. the students may not perceive these suffixes as psychologically real units. However, incorrect forms of these high-frequency derivatives produced by the students in the present study do show signs of awareness of word class affiliation. The suffix of choice by far when an incorrect derivative was produced for a noun-forming suffix is *-ness*, as in *ableness, *cautioness and *absentness. In addition to being one of the most frequent and productive suffixes in the English language, the suffix *-ness* is, since it is placed on a low difficulty level by Bauer and Nation, also highly regular and very predictable in character. Furthermore, as pointed out above, the suffix *-ness* carries high memory strength due to its relatively high parsing ratio, i.e. it is most often seen in derivatives that are decomposed. It may be that the students are intuitively aware of the usability of this suffix, i.e. put differently, it may mean that the students use the suffix *-ness* as a default suffix when being uncertain of the form of a noun derivative. However, in using *-ness* as a default suffix the students have disregarded the fact that there are morphological restrictions on the use of this suffix, namely that *-ness* can only be attached to adjective stems (Ljung, 2003: 92). Thus errors such as *borness, *entress and *achievness, where *-ness* has been added to verb stems, are all examples of overgeneralisation of this suffix. For the high-frequency word pairs discussed in the present paragraph, this error type appears to be most pervasive with *-ness,* but it can also be seen in connection with the derivative *couragable in which the suffix *-able* has been added to a noun stem even though it can only be attached to verb stems (Ljung, 2003: 102).

Although the students' understanding of word class affiliation is at its clearest with the suffix *-ness*, there are more errors that signal such awareness. For example, the incorrect forms *arriviation and *arrivement containing the noun-forming suffixes *-ion* and *-ment* were offered instead of the noun derivative *arrival* and *couragable and *courageful with the adjective-forming suffixes *-able* and *–ful* were produced instead of *courageous*. As a matter of fact, in this group few errors occurred for which the students showed ignorance of word class affiliation of the derivative forms. Instead the main problem seems to have to do with what type of stem the affix should be added to.

Finally, the students' high scores on the high-frequency word pairs ranging from *advise(d) – advice* to *acquaint(ed) – acquaintance* described above are also reflected in their own evaluation of the degree of knowledge they possess of these derivative forms (see Table 3). Not only are the majority of the students in most of these cases sure of the meanings of the words given (column KN(A)), but the majority of the students also indicate in most of these

cases that they are sure of what derivatives to form (column KN(B)). This applies to all but four derivatives (*ease/easiness, earliness, cleaner('s)* and *cleverness).* (Word pair 38 (*inordinate − inordinately*), as mentioned excluded from the discussion above, is an interesting anomaly since, even though 10 students were able to produce the correct derivative form, most of the students claimed that they did not know the word given nor that they were certain of the form of the derivative. It may be that the highly frequent and productive nature of the suffix *-ly* played a greater psychological role than the infrequency of the stem, thus helping the students produce the correct derivative.

The description offered of the twenty-one high-frequency word pairs discussed above stands in stark contrast to the word pairs on which the students scored the lowest, ranging from *contagious − contagion/contagiousness* (word pair 36) through *parsimonious − parsimony* (word pair 45). First of all, the great majority of these derivatives, with scores from 0 through 7, are not on West's general service list. In fact, most of these twenty word pairs are found among the last word pairs presented to the students in the present study, i.e. the majority of these words are highly infrequent even in the large BNC corpus. That the students were able to produce correct derivatives using, for example, the suffixes *-ness* and *-ity* in connection with the high-frequency items discussed above, but not with the highly infrequent items discussed in the present section (as in *surliness*, *voracity* and *asininity*), clearly shows the effect frequency has on the students' performance. The same can be seen with the suffix *-ous* with which most of the students were able to form the derivative *courageous* which forms part of a high-frequency word pair, but not the derivative *tempestuous* which is part of a low-frequency word pair. Furthermore, due to the infrequencies of most of these words, the relative simplicity of some of the intended suffixes seems irrelevant. Again the suffix *-ness* (Affix Level 3) can serve as an example in connection with the highly infrequent stem *surly* for which the majority of the students were not able to produce the correct derivative *surliness* (word pair 43). The suffix *-ful* to be removed from the word *guileful* to form *guile* (word pair 46) and belonging to Affix Level 4 may serve as a second example. Here too most of the students were unable to produce the correct form. Put simply, it seems that the infrequency of the word pairs takes its toll on the subconscious knowledge these learners may have of suffixation rules in English.

This contrast between the high-frequency word pairs discussed previously and the low-frequency word pairs discussed here is also clearly visible in the students' own evaluation of their knowledge of these word pairs. Whereas there is, as discussed above, a concentration of great certainty about the meanings of the words given and what words to be formed for the test items with scores from 15 through 10, most of the students indicate for the majority of these low-frequency word pairs their total lack of knowledge of both the words given (NK(A)) and the words to be formed (NK(B).

A typical trait of the errors produced in connection with this group of word pairs is that the students here appear to try a wider variety of suffixes. For instance, instead of the correct

derivative form *tempestuous* (word pair 37) students offered the incorrect forms **tempestish*, **tempestive*, **tempestant*, **tempested* and **tempesting* and instead of the correct noun derivative *parsimony* (word pair 45) the incorrect forms **parsimon, *parsimonia*, **parsimonionism* and **parsimonity* were seen. It seems that the more infrequent the stems are, the more uncertain the students are about what suffixes to use. Interestingly enough, in general the students here still seem to experience few problems with word class affiliation in that when, for example, an adjective is sought, as in the examples above, the students offer, though incorrect derivatives, stems to which another adjective-forming suffix has been added. Again, it is rather the semantic restrictions placed on the stems that cause the most problems for the students. **Exacerbateness*, **capacitable* and **tempestive* are all examples of errors where the students have added the right type of suffix to the wrong word class stem.

For the rest of the derivatives receiving average scores, from *annoy(ed) – annoyance* (word pair 32) through *bury* (*buried*) – *burial* (word pair 21), most of the students seem to be sure about the meaning of the word given, but less sure about the form of the derivative indicating in the majority of cases that they only think they know the form of the sought-for derivative, the solidifying of the students' knowledge of these derivatives thus depending on how many (additional) times they will encounter these derivative forms in the future. The uncertainty the students feel when dealing with this group of word pairs appears to have many reasons. In some cases, it may be the relatively high difficulty level of the suffix alone that may be troublesome for the students, as with the suffix *-ance* used to form the derivative *annoyance* (word pair 32) belonging to Affix Level 5, the stem *annoy* being a comparatively frequent item. In other cases, it may be the relative infrequencies of the stem and derivative in combination with a comparatively high difficulty level of the affix in question, such as with the derivative *burial* (the suffix *-al* also belonging to Affix Level 5) seen in word pair 21. In yet other cases, it may be the fact that the students were to remove the suffix rather than add one that made it difficult, as in word pair 15, where the students were to truncate the adjective *comfortable* and produce the noun *comfort*. The students here appear to be in a transitional stage in which they have gained knowledge of the meaning of the words given, but where a number of factors seem to hinder them from producing the correct derivatives.

No clear trends could be detected for the errors made by the students in connection with these word pairs, but the students here too seem to be intuitively aware of word class affiliation. Errors in which suffixes have been attached to the incorrect word class stem can however be seen.

As discussed in Section 2, studies show that for many L2 learners, it is very often the nature of the stem rather than the affix which determines whether a learner is able to produce a derivative form or not (see also Hancin-Bhatt & Nagy, 1994). The obvious problem with Bauer and Nation's framework then is that it totally ignores the role played by the stem. However, considering the nature of the stem alone also seems to be too simple an approach. Rather, the success or non-success with which a learner is able to produce a correct derivative

form appears, based on the students' results in the present study, to depend on an intricate interplay of three dimensions: the frequencies of the stem and derivative, the complexity of the suffix (complexity, as discussed above, based on the frequency, predictability, productivity and regularity of the suffix in question) and the relative frequencies of the stem and the suffix (in which the parsing ratio of the suffix is a useful tool). Thus it seems that if these three factors could be merged into a continuum, this continuum could help predict the ease or difficulty with which a learner will be able to produce a correct derivative form and this could, as an extension, help teachers when selecting items to be discussed in the classroom.

4.3.2. The results on the Swedish test

Whereas the English test result forms three main groups with clearly discernible differences, the corresponding Swedish test result can only be divided into two groups. The first one includes thirty-nine word pairs, starting with word pair 4 (*allvarlig – allvar* (*serious – seriousness*)) for which all 15 students gave the correct derivative down to, and including, word pair 36 (*hysterisk – hysteri* (*hysteric/al – hysteria*)) for which 11 students were able to produce the correct derivative form. Among these, high-frequency word pairs, as well as very infrequent word pairs, can be seen. It is of course due to their greater L1 vocabulary knowledge that the students are less susceptible to the frequencies of these word pairs than they are to the frequencies of the word pairs included in the English material even though, as pointed out before, the Swedish word pairs are less frequent than the ones included in the English test. The students are also less sensitive to the degree of productivity of the Swedish suffixes in that they do not only show mastery of highly productive suffixes, such as the verb-forming suffix *-era* (as in *verifiera* (*verify*) (word pair 46), but also of suffixes that only show some productivity, such as the noun-forming suffix *-het* (as in *vilsenhet* (*forlornness*) (word pair 37) and of those suffixes that are not productive at all, such as the noun-forming suffix *-dom* (as in *fattigdom* (*poverty*) (word pair 20) (for productivity of Swedish suffixes see Liljestrand (1993) and Möijer (1998)). Furthermore, for all of these items, most of the students indicated that they were not only sure about the meaning given, but also about the derived form sought. Thus this group clearly corresponds to the group of high-frequency word pairs ranging from *advise(d) – advice* through *acquaint(ed) – acquaintance* seen in the English material. Based on the few errors in connection with this group of derivatives, it is impossible to decide whether the students actually perceive the suffixes as psychologically real units or whether the derivatives have been acquired as unseparated wholes. The students do not, however, appear to have used any suffix as a default suffix, which seemed to be the case with the suffix *-ness* in the English test, especially with the high-frequency word pairs.

For the majority of the remaining word pairs (7 out of 11), having scores from 4 through 11, the students are still sure of the meaning of the word given, but less sure of the word to be

formed, thus being in a transitional stage. This group, again like the corresponding English group, incorporates both high- and low-frequency word pairs for which the students experience difficulties in producing the correct derivative for a variety of reasons. For example, whereas the word given in word pair 8 (*aktuell(t) (of current interest)*) is highly frequent making up the high frequency of the word pair almost entirely on its own, the word sought for (*aktualisera (bring something to the fore)*) is highly infrequent. This means that whereas the students probably have come across the adjective *aktuell(t)* quite a few times, they have most likely not dealt with the corresponding verb as many times, thus rendering *aktualisera* a difficult word to produce even though the suffix itself is highly productive. (The same situation could be seen with word pair 32 (*annoy – annoyance*) which formed part of the corresponding miscellaneous group in the English material.) The derivative *gondoljär* (*gondolier*) forming part of word pair 50 is difficult for other reasons. This word pair does not only involve a stem and a derivative which are both very infrequent words, but the suffix itself (*-jär)* is neither frequently used nor productive. The same complexity was noticed in connection with word pair 21 (*bury (buried) – burial*) also belonging to the corresponding group in the English test. Consequently, it is with this miscellaneous group of both high- and low-frequency word pairs that the complex interplay noticed and discussed in connection with the English material is most pronounced in the Swedish suffixation test.

4.4 Individual L1 and L2 suffixation knowledge compared and the correlation between suffixation knowledge and vocabulary size

Tables 5 and 6 present the students' individual scores on the English and Swedish test respectively.

STUDENT	SCORE FOR **PART A** **ENGLISH TEST** (50 ITEMS)
STUDENT 5	36
STUDENT 15	35
STUDENT 11	34
STUDENT 2	33
STUDENT 3	32
STUDENT 7	31
STUDENT 14	
STUDENT 4	27
STUDENT 9	
STUDENT 1	26
STUDENT 8	25
STUDENT 12	22
STUDENT 6	20
STUDENT 13	
STUDENT 10	15

Table 5. The students' individual scores on the English test

STUDENT	SCORE FOR **PART A** **SWEDISH TEST** (50 ITEMS)
STUDENT 11	50
STUDENT 5	47
STUDENT 6	
STUDENT 9	
STUDENT 3	46
STUDENT 12	
STUDENT 2	44
STUDENT 8	
STUDENT 7	43
STUDENT 14	
STUDENT 4	41
STUDENT 13	39
STUDENT 1	38
STUDENT 15	
STUDENT 10	31

Table 6. The students' individual scores on the Swedish test

These results show that for 10 of the 15 students, there is a correlation between their ability/inability to form derivatives in their first language as compared to their second language. While 6 students consistently obtained scores higher than the mean in both languages (highlighted in bold), 4 students performed poorly in both languages (indicated in italics), achieving below the mean in both languages (mean Eng: 27.6; mean Swe: 42.9). This transfer of affixation skills from a learner's L1 to his/her L2 has been noticed in other studies too (e.g. Piasecka, 2006: 247). As for the other 5 students, no clear patterns could be discerned.

Furthermore, not only is there a correlation between the students' ability/inability to form derivatives in their native language and their L2, but also between their suffixation skills and the number of words they know (see also Schmitt and Meara (1997) for the same observation). In Karlsson (2012), the same 15 students involved in the present study were tested on their general vocabulary knowledge (such as vocabulary taught at upper secondary school level, advanced vocabulary, specialized vocabulary, idiomatic expressions and proverbs, phrasal verbs, polysemous words, near synonyms and false friends) in both their L1 and L2. The students who did well on those tests did generally also well on the suffixation tests; those students who did poorly, normally did so on both types of tests.

5. CONCLUSIONS AND PEDAGOGICAL IMPLICATIONS OF THE RESULTS

From a quantitative perspective, the students performed, as expected, considerably better on the L1 test (mean: 42.9, i.e. 85.8% correct answers) than on the L2 test (mean: 27.6, i.e. 55.2% correct answers). Furthermore, even though the students' individual results varied a great deal on both tests (see Tables 5 and 6), the achievements show that a learner's L1 knowledge tallies with his/her L2 mastery to a great extent. In fact, for as many as ten of the

15 students the L1 and L2 results showed similarities, five learners being high-achievers and five being low-achievers in both languages.

Furthermore, overriding all other aspects of the students' results on the individual words is the role played by the frequencies of the word pairs, which suggests that word family is indeed a psychologically justified concept. Based on this observation, the learners' achievements on the English test could be divided into three areas with different characteristics. Firstly, there are those high-frequency items (as seen in the BNC and also substantiated by West's service list (1953)) of which the learners displayed a solid knowledge of the word given as well as the derivative sought. Here the complexity of the suffix (in terms of Bauer and Nation's approach presented in Subsection 2.2) did not seem to matter at all. These are thus words that the students must have encountered quite a few times during their English studies and may therefore approach as unseparated wholes. Slightly more than 40% of the word pairs were seen to belong to this type. Interestingly, on the relatively few occasions the students provided the incorrect derivative form for these test items, *-ness*, a suffix placed on a low difficulty level by Bauer and Nation, was used as default suffix. On the Swedish test, almost 80% of the word pairs could be described in the same way as the English test items above (except for there not being an apparent suffix used as a default suffix). Also, similar to the results on the English test, high- and low-frequency items as well as those showing high and low productivity could be seen in this group.

Secondly, at the other end of the spectrum, there are those low-frequency items (both word given and word to be produced (again substantiated by the BNC and West's service list) that very few of the students knew. One typical characteristic of these test items is that even when the learners were to add less complex suffixes which they mastered in connection with the high-frequency items discussed above, they failed. Also, no default suffix could be detected in this category, rather a wide variety of incorrect suffixes were used for one and the same derivative. While around 30% of the word pairs adhere to this description on the English test, none of the Swedish word pairs were of this type, showing the higher level of mastery of the learners' mother tongue.

Thirdly, the most interesting group is the one that falls between the two described above. For slightly more than 15% of the English word pairs (and almost as many of the Swedish word pairs), the students here most often indicated that they knew the stem but that they experienced uncertainty about how to form the derivative. Put differently, in this group the students' knowledge of the stems seems to have been solidified, while they are still in the process of developing an understanding of the make-up of that stem's word family. Furthermore, based on the results in this category, conclusions could be drawn that the degree of success with which a learner is able to produce a correct derivate form depends on a convoluted interplay among the frequencies of the stem and derivative, the complexity of the suffix (where complexity is based on the frequency, predictability, productivity and regularity of the suffix in question) and the relative frequencies of the stem and the suffix.

Despite the differences between the three result types described above, all the errors made by the students show two similarities. The suffix chosen, even when it was incorrect, still most often displayed the sought-for word class affiliation, indicating the students' awareness of the meaning of the suffixes. Rather, when errors occurred, the students seemed to fail to understand that there are restrictions as to what stems certain suffixes can be attached to.

Finally, during informal discussions with the students after they had taken the two suffixation tests, it came to the fore that hardly any of them had ever worked with affixation in school before, in either language. On a more general note, since the construct of word family does not only appear to be a linguistic term used for a set of words the meanings of which have close affinities, but, as many studies have shown including the present one, a psychologically justified concept, students should, when taught a specific word, also be confronted with its derivative forms to a greater extent than what apparently is the case and in connection with these forms be taught what the affixes in question mean and do and to what stems they can be attached. As a matter of fact, a growing body of research indicates that explicit training indeed helps learners (L1 as well as L2 learners) to transfer their receptive affixation knowledge into productive use (for L1 children, see Cunningham (1998), and for L2 adults, see Mirhassani and Toosi (2000)).

Furthermore, the present study shows that for quite a few students there is a clear correlation between what they know and can do in terms of affixation in their L1 as compared to their L2. It therefore seems that affixation knowledge gained in one's native language may be transferred to an L2 (and vice versa?), thus automatically increasing the student's vocabulary knowledge in two or more languages at the same time. This, if anything, should be an incentive to introducing discussions and exercises on affixation in the language classroom. Teachers should, however, be aware that 'word family' is not an equally real concept to all learners. Instead, the validity of this construct primarily appears to depend on the maturity level of the learner in question (Gardner, 2007) and discussions and instruction should be shaped accordingly. For example, teachers would probably gain the most, especially with weak learners, by starting with derivatives that are regular in character (i.e. those that involve few changes in pronunciation, spelling etc.) and not until these have been mastered move on to more irregular ones. Recent research supports such an approach. In Gor and Jackson (2013), for instance, it is demonstrated that L1 as well as L2 learners initially cognitively prefer to deal with inflectional suffixes that cause no allomorphy to slowly moving on to more and more complex stem allomorphy. It is probably not too far-fetched to assume this to be the case with derivational suffixes too.

NOTES

1. *Complacenc*y (1) and *complacement* (2) are also considered to be correct answers. They are listed here for spatial reasons.
2. The derivative *glömskhet* is also regarded to be correct.
3. Other forms judged to be acceptable are *säsongartat* (1), *säsongsaktigt* (1), *säsongsberoende* (3) and *säsongsmässigt* (1).

REFERENCES

Anderson, R. C. (1984)'How many words are there in printed school English? *Reading Research Quarterly, 19*, 304–330.

Andersson, L-G. (1987) *Språktypologi och språksläktskap*. (3rd ed.). Stockholm: Skriptor/Almqvist Wiksell.

Bauer, L. & Nation, P. (1993) Word Families. *International Journal of Lexicography, 6*(4), 253–279. Oxford: Oxford University Press.

Berko, J. (1958). The child's learning of English morphology. *Word, 14*, 150–177.

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

Bird, N. (1990) *A First Handbook of the Roots of English*. Hong Kong: Lapine Education and Language Services Ltd.

Bradley, D. C. (1980). Lexical representation of derivational relation. In M. Aronoff & M. L. Kean (Eds.), *Juncture: A Collection of Original Papers* (pp. 35–55). Saratoga, CA: Anma Libri.

Caramazza, A., Laudanna, A. & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition, 28*, 297–332.

Carlisle, J. F. (2000) Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing: An Interdisciplinary Journal, 12*, 169–190.

Cunningham, P. (1998). The multisyllabic word dilemma: Helping students build meaning, spell, and read "big" words.' *Reading and Writing Quarterly: Overcoming Learning Difficulties, 14*, 189–218.

De Jong, N. H., Schreuder, R. & Baayen, R. H. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes, 15*(4/5), 329–365.

Derwing, B. L. & Baker, W. J. (1979). Recent research on the acquisition of English morphology. In Fletcher, P. & Garman, M. (Eds.), *Language acquisition* (pp. 209–223). Cambridge: Cambridge University Press.

Gardner, D. (2007) Validating the construct of word in applied corpus-based vocabulary research: A critical survey. *Applied Linguistics, 28*(2), pp. 241-265.

Gor, K. & Jackson, S. (2013) Morphological decomposition and lexical access in a native and second language: A nesting doll effect. *Language and Cognitive Processes, 28*(7), 1065–1091.

Hancin-Bhatt, B. & Nagy, W. (1994) Lexical transfer and second language morphological development. *Applied Psycholinguistics 15*, 289–310.

Haspelmath, M. & Sims, A. D. (2010) *Understanding morphology*. (2nd ed.). London: Hodder Education.

Karlsson, M. (2012). *Quantitative and Qualitative Aspects of Advanced Students' L1 (Swedish) and L2 (English) Knowledge of Vocabulary*. Halmstad: Högskolan i Halmstad.

Liljestrand, B. (1993) *Så Bildas Orden: Handbok i Ordbildning*. Lund: Studentlitteratur.

Ljung, M. (2003). *Making Words in English*. Lund: Studentlitteratur.

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

Mirhassani, A. & Toosi, A. (2000). The impact of word-formation knowledge on reading comprehension. *IRAL, 38*, 301–311.

Mulder, K., Schreuder, R. & Dijkstra. (2013) Morphological family size effects in L1 and L2 processing: An electrophysiological study. *Language and Cognitive Processes 28*(7), 1004–

1035.

Möijer, K. (1998) *Svensk Ordlära: Så Bildas Orden.* Solna: Ekelund.

Nagy, W. E., Anderson, R. C., Schommer, M., Scott, J. A. & Stallman, A. (1989). Morphological families in the internal lexicon. *Reading Research Quarterly 24*, 263–282.

Nation, I. S. P. (2001) *Learning Vocabulary in another Language.* Cambridge: Cambridge University Press.

Piasecka, L. (2006) 'Don't lose your head' or how Polish learners of English cope with L2 idiomatic expressions. In J. Arabski (Ed.), *Cross-linguistic Influences in the Second Language Lexicon* (pp. 246–258). Buffalo: Multilingual Matters.

Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning, 48*(2), 281–317.

Schmitt, N. (1999). The relationship between TOEFL vocabulary items and meaning, association, collocation, and word class knowledge. *Language Testing, 16*, 189–216.

Schmitt, N. & McCarthy, M. (Eds.) (1997). *Vocabulary: Description, Acquisition and Pedagogy.* Cambridge: Cambridge University Press.

Schmitt, N. & Meara, P. (1997). Research vocabulary through a word knowledge framework – word associations and verbal suffixes. *Studies in Second Language Acquisition, 19*, 17–36.

Schmitt, N. & Zimmerman, C. B. (2002). Derivatives word forms: What do learners know? *TESOL Quarterly, 36*(2),145–171.

Schreuder, R. & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language, 3*, 118–139.

Smedts, W. (1988). De beheersing van de nederlandse woordvorming tussen 7 en 17. In van Besien, F. (Ed.), *First Language Acquisition. ABLA Papers 12* (pp. 103–127). Antwerp: Association Belge de Linguistique Appliquée/Universitaire Instelling Antwerpen.

Thorndike, E. L. (1941) *The Teaching of English Suffixes.* Columbia University: Teachers College.

Tyler, A. & Nagy, W. (1989). The acquisition of English derivational morphology. *Journal of Memory and Language, 28*, 649–667.

Tyler, A. & Nagy, W. (1990). Use of derivational morphology during reading. *Cognition, 36*, 17–34.

West, M. (1953). *A General Service List of English Words.* London: Longman.

White, T. G., Power, M. A. & White, S. (1989). Morphological analysis: Implications for teaching and understanding vocabulary growth. *Reading Research Quarterly, 24*, 283–304.

Wysocki, K. & Jenkins, J. (1987). Deriving word meanings through morphological generalizations. *Reading Research Quarterly, 22*, 66–81.

APPENDIX: TRANSLATION OF TABLE 4

No.	WORD GIVEN A	WORD TO BE FORMED B
4	serious	seriousness 15
19	bicycle	cyclist 15
18	forgot (forget)	forgetfulness 15
24	full (re: food)	satisfaction (re: food) 15
30	brutal	brutality 15
43	tolerate	tolerant 15
2	week	week by week 15
9	(the) organisation	organise 15
17	arrange(d)	arrangement 15
16	journalist	journalism 15
21	(the) theory	theoretic(al) 15
25	shame	ashamed 15
29	shy	shyness 15
33	boast(s)	boastful 15
1	suggestion	as a suggestion 15
37	forlorn	forlornness 15
3	in the very centre	centre 15
13	acceptable	accept 15
20	poor	poverty 15
11	offer(s) (verb)	offer (noun) 15
42	simple-minded	simple-mindedness 15
7	obvious	obviously 14
39	fall in love	love (affair) 14
34	idiot	idiocy 14
31	(the) comfort	comfortable 14
35	statue	statuette 14
15	dance	female dancer 13
48	confirm (in church)	confirmand 13
5	drive (driven)	fit to drive 13
46	verification	verify 13
38	theology	theologian 13
32	tolerate	hardy 12
14	aware (adjective)	awareness 12
22	doubt(s) (noun)	of doubtful reputation 12
45	stand-up	stand-up comedian 12
27	scholarship	holder of a scholarship 12
40	fanaticism	fanatic 12
28	(highest) fever (measured)	free from fever 12
36	hysteric/al	hysteria 11
47	instruct	instruction 11
23	(the) truth	truthful 11
8	of current interest	bring something to the fore 11
26	lion	lioness 10
12	(has) cost	costly 10
41	(the) quantity	quantitative 10
10	(peak) season	seasonal 10
49	fabrication	fabricate 9
50	gondola	gondolier 7
44	in an isolated place	recluse 6
6	(the) agreement	according to agreement 4