Linguistic Research 28(1), 95-115

Ordering Paradoxes in English Suffixation: An Account Based on Relative Frequency*

Ju-Eun Lee · Sung-Hoon Hong*****
(Hankuk University of Foreign Studies)

Lee, Ju-Eun & Hong, Sung-Hoon. 2011. Ordering Paradoxes in English Suffixation: An Account Based on Relative Frequency. *Linguistic Research* 28(1), 95-115. This paper is an attempt to explain ordering paradoxes found in English suffixation. On the assumption that the order of suffix attachment follows Siegel's (1974) Level Ordering Hypothesis, we obtain attested suffix combinations from the CELEX lexical database (Baayen et al. 1995), and analyze them using the concept of productivity proposed by Hay & Baayen (2002). It is found, however, that the productivity approach alone does not explain the attested order of suffixes properly. In this paper, we rather propose an account of ordering paradoxes based on the notion of relative frequencies (Hay 2002, 2003): Ordering paradoxes come about when the relative frequency of a suffixed word is greater than 1, or put it differently, when a suffixed word is used more frequently than its base word so that the suffixed word behaves like one lexicalized whole. **(Hankuk University of Foreign Studies)**

Key Words suffix order, Lever Ordering Hypothesis, ordering paradoxes, productivity, relative frequencies

1. Introduction

It has been known that suffixation does not occur arbitrarily but is usually governed by certain rules. One of the well-known model for the rules of suffixation is Siegel's (1974) Level Ordering Hypothesis (LOH), where she proposes that morphological rules are intermingled systematically with phonological rules, and that the suffixes are divided into two classes, Level I and II suffixes, of which the first comes closer to the base than the second. As noted by many researchers (Selkirk 1982, Strauss 1982, Kiparsky 1982, Mohanan 1986, Fabb 1988, Spencer 1991, Hay

* This work was supported by Hankuk University of Foreign Studies Research Fund of 2010. We are sincerely grateful to anonymous reviews for their helpful comments.
** First author.
*** Corresponding author.

2002), however, not all suffix combinations follow this order, and Level II suffixes may come before Level I suffixes.

The purposes of this paper is to explain such cases of "ordering paradoxes," which do not follow the general suffix order as posited by Siegel's LOH. For the purpose of this paper, we will deal in particular with the following aspects of the question: how suffix orders are displayed in actually derived words, whether there is any correlation between leveled suffixes and productivity, and what a plausible alternative account of ordering paradoxes is.

The organization of this paper is as follows: Section 2 illustrates the possible and existing suffix combinations on the basis of LOH. To investigate the order of suffixes, existing suffix combinations are extracted from the CELEX lexical database (Baayen et al. 1995). We then examine these combinations to identify five cases of ill-formed combinations which do not follow the general suffix order. In section 3, the procedure of computing productivity is introduced to investigate the possibility of productivity being the determining factor for suffix orders. In section 4, Hay's relative frequency is introduced and adopted for a possible solution to those problematic suffix combinations. Finally, section 5 presents the conclusions that can be drawn from this paper.

2. LOH and Ordering Paradoxes

2.1 LOH and the List of Level I & II Suffixes

English suffixation is constrained fundamentally by subcategorization in such a way that a certain suffix attaches only to a certain syntactic category. What is more important, however, is the fact that the internal structure of a word is not determined only by the selectional restrictions, but also by the order of suffixes. This constraint reduces the possible set of morphologically complex words, and explains why a word is derived with a certain suffix beyond categorical selectional restrictions.

According to Siegel's LOH, suffixes are classified into two groups, Level I and II suffixes, associated respectively with a morpheme boundary and a word boundary. Level I and Level II suffixes can be attached recursively in each level, but the order between two groups must be that in which Level I suffixes come before Level II

suffixes.

In this paper, we examine 26 suffixes out of 89 suffixes in the Marchand's (1969) list of English suffixes. We do not consider suffixes showing allomorphy (e.g. *-ance~ence*, *-ancy~ency*, *-ant~ent*), suffixes that are not in Hay & Plag's (2004) list (e.g. *-ard*, *-ure*, *-wise*), suffixes that carry their own stress (e.g. *-ee*, *-eer*, *-ette*), and suffixes that do not show clear morpheme boundaries. Here is the list of 26 suffixes:

(1) -able, -age, -al, -an, -ate, -ary, -ion, -don, -ful, -hood, -ic, -ify, -ish, -ism, -ist, -ity, -ive, -ize, -less, -ly, -ment, -ness, -ory, -ous, -ship, -some

These 26 suffixes are then divided into two classes, Levels I and II, based on whether they trigger phonological changes such as stress shift and segmental changes (more specifically, Trisyllabic Shortening, Spirantization, and Velar Softening (cf. Chomsky & Halle 1968)). If a derived word undergoes one of these phonological changes, the suffix is classified as Level I; if not, it is grouped as Level II. Following is the list of Level I and II suffixes that we adopt in this paper:

(2) a. Level I suffixes[1]:
-al$_{(adj)}$, -an, -ary, -ate, -ic, -ify, -ion, -ist, -ity, -ive, -ism, -ize, -ory, -ous
b. Level II suffixes[2]:
-able, -age, -dom, -ful, -hood, -ish, -less, -ly, -ment, -ness, -ship, -some

---

[1] Suffixes such as *-ism*, *-ist* and *-ize*, have been identified as dual membership suffixes (cf. Aronoff 1976, Selkirk 1982, Aronoff & Sridhar 1983, Szpyra 1989, Giegerich 1999). They are classified as Level I suffixes here because they can trigger segmental changes such as Velar Softening (e.g. *lyricist, analogize*) and Trisyllabic Shortening (e.g. *patronize, divinity*).

[2] Although it changes the position of the primary stress in some examples, *-age* is categorized here as Level II because it is a transparent suffix that does not change the stress of a word in general. In fact, we found that there are only two examples of stress change, *equíp* → *équipage* and *cóncubine* → *concúbinage*, among the 113 *-age* derivatives in the CELEX lexical database. *-able* is also classified as a Level II suffix although it was treated as a dual membership suffix in some literature (Aronoff 1976, Selkirk 1982, Szpyra 1989, and Anderson 1992). This is because *-able* exhibits dual behaviors only in *-able+ity* combination. In this paper, we posit that *-able+ity* is restructured as one morphological unit so that it behaves like one suffix which we treat as a Level II suffix.

2.2 Attested Suffix Orders and Ordering Paradoxes

One of the most significant issues to be addressed regarding suffixation is ordering paradoxes. As mentioned previously, Level II suffixes in general come after Level I suffixes, but not vice versa. However, there are some suffix combinations which do not follow this order.

Let us first examine the existing suffix combinations in terms of the levels to which each suffix belongs. For this purpose, we draw on the CELEX lexical database, as it shows the details of the internal structures of complex words. It is found that there are 110 suffix combinations from CELEX, whose skeletal structures are given below:[3]

(3) Skeletal structures of existing combinations

Skeletal Structure (the number of the combination found)	Examples
a. Base + I*[4] (55)	nation-al$_{I}$-ity$_{I}$, convent-ion$_{I}$-al$_{I}$-ity$_{I}$, institute-ion$_{I}$-al$_{I}$-ize$_{I}$-ate$_{I}$-ion$_{I}$
b. Base + II* (13)	care-less$_{II}$-ness$_{II}$
c. Base + I* + II* (33)	ruin-ous$_{I}$-ly$_{II}$, symmetr-ic$_{I}$-al$_{I}$-ness$_{II}$, emot-ion$_{I}$-less$_{II}$-ly$_{II}$
d. Base + II + I* (6)	ramp-age$_{II}$-ous$_{I}$, govern-ment$_{II}$-al$_{I}$, argue-ment$_{II}$-ate$_{I}$-ion$_{I}$ complete-ment$_{II}$-ary$_{I}$, argue-ment$_{II}$-ate$_{I}$-ive$_{I}$, environ-ment$_{II}$-al$_{I}$-ist$_{I}$
e. Base + II + I* + II (3)	ramp-age$_{II}$-ous$_{I}$-ly$_{II}$, environ-ment$_{II}$-al$_{I}$-ly$_{II}$, argue-ment$_{II}$-ate$_{I}$-ive$_{I}$-ly$_{II}$

The skeletal structures of English derived words are divided into five types. In

[3] See Appendix 1 for the exhaustive list of 110 suffix combinations.

[4] The asterisk indicates that there can be more than one appearance of the suffix.

(3a) and (3b), suffixes of one level attach to the bases. In (3c), there appear two types of suffixes, Level I and II suffixes, which are combined in this order. In this type, a maximum of three Level I suffixes is followed by one Level II suffix. The last two types, (3d) and (3e), exhibit the cases of ordering paradox where a Level II suffix is placed before a Level I suffix. Among the existing 110 combinations, nine cases of ordering paradox are found, all of which are listed in the table above.

If we look at the cases of ordering paradox closely, we can see that there are two Level II suffixes that come before a Level I suffix: *-age* and *-ment*. These suffixes appear in a particular combination with a Level I suffix as in *-age*$_{II}$*ous*$_{I}$, *-ment*$_{II}$*al*$_{I}$, *-ment*$_{II}$*at*$_{I}$*ion*$_{I}$, *-ment*$_{II}$*ary*$_{I}$, *-ment*$_{II}$*at*$_{I}$*ive*$_{I}$, and *-ment*$_{II}$*al*$_{I}$*ist*$_{I}$ (Base+II+I*); *-age*$_{II}$*ous*$_{I}$*ly*$_{II}$, *-ment*$_{II}$*al*$_{I}$*ly*$_{II}$, and *-ment*$_{II}$*at*$_{I}$*ive*$_{I}$*ly*$_{II}$ (Base+II+I*+II). These are nine apparent cases of ordering paradox, but some of the suffix combinations overlap, *-age*$_{II}$*ous*$_{I}$ and *-ment*$_{II}$*al*$_{I}$. Thus, we can collapse ordering paradoxes into five types, *-age*$_{II}$*ous*$_{I}$, *-ment*$_{II}$*al*$_{I}$, *-ment*$_{II}$*at*$_{I}$*ion*$_{I}$, *-ment*$_{II}$*ary*$_{I}$, and *-ment*$_{II}$*at*$_{I}$*ive*$_{I}$. These are the ordering paradoxes for which we will seek an explanation in this paper.

So far, we have examined the fundamental structures of suffix combinations attested in CELEX. We have found five types of suffix combinations that do not follow the general suffix order. In the next section, we will consider the notion of productivity as an account of suffix orders and ordering paradoxes.

3. Productivity-based Approach to Suffix Order

Hay & Baayen (2002) discuss the premise that suffix attachment is related to the productivity of each suffix. According to this productivity approach, suffixes with higher productivity attach outside suffixes that have lower productivity. For instance, *-ism* with the productivity 0.005 attaches outside the suffix *-ion*, which has the productivity 0.001, as in *expressionism*. The goals of this section are to investigate whether suffix orders can be explained in terms of productivity and whether there is any correlation between productivity and the suffix levels. For these purposes, we will introduce a formula for measuring productivity provided by Baayen (1989) and the productive value of 26 leveled suffixes.

The productivity of suffixes depends on how frequently suffixes are used and how many times *hapaxes* appear in a corpus. A hapax, which is an abbreviated form

of *hapax legomena*, indicates a word that occurs only once in a surveyed data set. The formula suggested by Baayen (1989) for calculating productivity is as follows:

(4) P = n1 / N (P: productivity, n1: hapaxes, N: token frequency)

The productivity value of a certain suffix *P* is calculated by dividing the total frequency number of hapaxes, *n1*, by the total frequency number of the derived words containing the suffix *P*, i.e. *N*. An example of a trial computation is performed as follows:

(5) baby*hood* (73) kitten*hood* (11) planet*hood* (9) monk*hood* (6)
self*hood* (1) prince*hood* (1) daddy*hood* (1) ego*hood* (1)

In the sample data set above, the token frequency of the words containing the suffix *-hood* is 103 (73+11+9+6+1+1+1+1), and there are four occurrences of hapaxes. When we input these figures in the formula (4), we can get the productivity of *-hood*, which is 4/103=0.003.

Using this formula, Hay & Baayen (2002) calculated the productivity values of 80 affixes, 26 prefixes and 54 suffixes. In this paper, however, we do not adopt their results for the following reasons: First, the data size set from the CELEX lexical database on which their calculation is based is not big enough to reflect the real language usage. Second and more significantly, their productivity values were shown only to three decimal points, which poses a problem because productivity values calculated by (4) are usually so small that they cannot be differentiated from one another by using only three decimal points. For these reasons, the productivity of each suffix is recalculated in this paper based on the frequency information obtained from the Corpus of Contemporary American English (COCA; Davies 2008), which has a more extensive data set (about 400 million words) than CELEX.

The newly calculated productivity values for the 26 suffixes are presented below in Table 1 (for Level I suffixes) and Table 2 (for Level II suffixes).[5]

---

[5] Tables 1 and 2 show that there are noticeable differences in the productivity values of suffixes between our study and Hay & Baayen (2002). The productivity value of *-ship* in our study, for instance, is calculated as 0.0032, while Hay & Baayen's value is 0.009. The difference arises presumably due to the size of the data set from which productivity is obtained.

Table 1. Level I suffixes and the productivity values

Suffix	Type F	Token F	Hapaxes	Productivity	Hay & Baayen's P
-ion	12166	4145549	5888	0.0014	0.001
-ist	2420	430000	718	0.0017	0.005
-ate	2004	475671	987	0.0021	0.003
-ify	295	65160	141	0.0022	0.002
-ory	1963	483540	1129	0.0023	0.002
-ary	1576	373149	899	0.0024	0.001
-ity	6969	1520224	3832	0.0025	0.001
-ive	6437	1486031	3841	0.0026	0.003
-al$_{(adj)}$	6340	879896	2767	0.0031	0.001
-ic	7457	988780	3418	0.0035	0.002
-ous	3995	650669	2280	0.0035	0.001
-ize	1066	128329	458	0.0036	0.001
-an	6692	722708	3379	0.0047	0.003
-ism	2523	201673	1106	0.0055	0.003

Type F(requency): Number of distinct words which contain the suffix.
Token F(requency): Total number of the derived words.

Table 2. Level II suffixes and the productivity values

Suffix	Type F	Token F	Hapaxes	Productivity	Hay & Baayen's P
-ment	1780	1285556	822	0.0006	0.000
-ful	1284	335550	799	0.0024	0.002
-age	1907	356037	1115	0.0031	0.002
-ship	1056	193758	626	0.0032	0.009
-hood	453	72761	280	0.0038	0.004
-less	1572	183552	734	0.0040	0.017
-ly	2634	362572	1689	0.0047	0.001
-dom	529	75281	363	0.0048	0.002
-able	4832	497259	2675	0.0054	0.003
-ish	3390	371761	2040	0.0055	0.005
-some	318	23042	142	0.0062	0.009
-ness	5779	431894	2778	0.0064	0.008

Since it is generally assumed that Level II suffixes are more productive than Level I suffixes, the productivity values for Level II suffixes are predicted to be higher than those for Level I suffixes. Further, since the general suffix order is that in which Level I suffixes precede Level II suffixes, it is also predicted that suffixes that come closer to a base is lower in the productivity value than those that come farther from a base.

We found that both predictions were not borne out, however. As we see in Tables 1 and 2, the ranges of the productivity values for Level I and II suffixes overlap with each other (0.0014~0.0055 and 0.0006~0.0064, respectively). This overlap made it difficult to analyze suffix order based on productivity values[6]. Moreover, we examined the 110 existing suffix combinations in terms of the suffix productivity values, but it was turned out that 39 combinations did not exhibit the predicted productivity order. In these combinations, which are given below, a high productivity suffix precedes a lower productive suffix. (The portion of suffix combination that does not follow the expected productivity order is marked within brackets.)

(6) a. [able+ity], [al+ist], [al+ist]+ic, ion+[al+ist], [al+ity], ion+[al+ity], [an+ity], ity+[an+ic+al], [ate+ion], [ate+ion]+al, [ate+ion]+ism, [ate+ion]+ist, [ic+al], ist+[ic+al], [ic+ate], [ic+ize], [ic+ate+ion], [ify+ate], [ify+ion], [ity+ary], [ive+ist], [ive+ate+ion], [ize+ion], al+[ize+ion], ion+[able+ity], ion+al+[ize+ion], ive+[ize+ate+ion], [ous+ity]

b. [able+ly], [ish+ly], [ly+hood], [some+ly]

c. ify+[able+ly], ion+[able+ly], [an+ship], ity+[an+ic+al]+ly, [ic+al]+ly, [ic+al]+ness, [ize+ment]

d. ment-[ate-ion], ment-[al-ist]

Note that exceptions to the productivity order are not just found in regular cases of suffix order that conform to LOH as in (6a), (6b), and (6c) (Base+I*, Base+II*, and Base+I*+II*, respectively) but also in cases of ordering paradoxes (6d)

[6] In fact, the difference between the productivity values of Level I and II suffixes is found to be statistically significant (t=-2.225, p=0.036 (two-tailed), df=24). This only shows that productivity values can distinguish Level I and II suffixes in general, but still it does not provide a proper account of suffix order.

(Base+II+I*), implicating that productivity cannot be an account of ordering paradox, either.

So far, we have seen that the attested suffix orders do not follow as predicted by the productivity theory, and hence, productivity alone cannot provide a proper account of suffix order and ordering paradoxes. In the following section, adopting LOH as a theory for the general suffix order, we will seek an account of ordering paradox based on the notion of relative frequency proposed by Hay (2002, 2003).

4. Ordering Paradoxes in Terms of Relative Frequency

We have found earlier that there are Level II suffixes which may come before Level I suffixes: *-age* and *-ment*. These suffixes appear with Level I suffixes only in combinations such as *-age*$_{II}$*ous*$_{I}$, *-ment*$_{II}$*al*$_{I}$, *-ment*$_{II}$*at*$_{I}$*ion*$_{I}$, *-ment*$_{II}$*ary*$_{I}$, *-ment*$_{II}$*at*$_{I}$*ive*$_{I}$, *-ment*$_{II}$*al*$_{I}$*ist*$_{I}$, *-age*$_{II}$*ous*$_{I}$*ly*$_{II}$, *-ment*$_{II}$*al*$_{I}$*ly*$_{II}$, and *-ment*$_{II}$*at*$_{I}$*ive*$_{I}$*ly*$_{II}$. The last four combinations include one of the ill-formed suffix combinations, *-age*$_{II}$*ous*$_{I}$, *-ment*$_{II}$*al*$_{I}$, and *-ment*$_{II}$*ative*$_{I}$. Excluding these four combinations, we are left with five suffix combinations of the skeletal structure type Base+II+I*. In what follows, our discussion of ordering paradox will thus be limited to these five cases.

4.1 Relative Frequency

Hay (2002, 2003) demonstrates the parsability of a complex word based on a dual route model. For every suffixed word, the perception process of a complex word is divided in two ways. One is a decomposed access route, and the other is a whole-word access route, as depicted in the following figure:

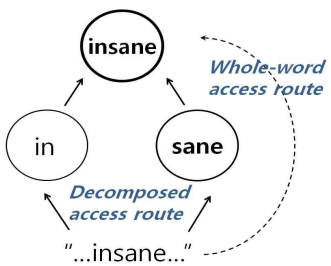

Figure 1. Schematized dual route model (Hay 2003, p.11)

The figure shows how the structure of a word can be perceived. When an affix can be easily separated from its base, the word is accessed through the decomposed access route. In this case, speakers easily notice that the word is composed of its subparts. This route is activated when the base word is more frequently used than the derived word. In contrast, when an affix and its base are tightly knit and cannot be broken apart easily, the whole-word access route has the advantage. This happens when the frequency of a derived word is higher than that of its base. The formula for calculating relative frequency, proposed by Plag (2002: 10) based on Hay (2002), is provided below:

(7) $f_{relative} = f_{derivative} / f_{base}$

$f_{relative}$: relative frequency

$f_{derivative}$: summed frequency of the derived word

f_{base}: summed frequency of the base word

According to this formula, relative frequency is greater than 1 when a derived word is used more frequently than its base. This implies that this word should be considered one lexical word rather than a word composed of a base and an affix, and that this word is processed through the whole-word access route. In contrast, relative frequency is smaller than 1 when a base word is more frequent than the derived word. In this case, the derived word is processed through the decomposed access route, and thus can be broken down into a base and an affix.

Suffixes, therefore, are categorized into two types: suffixes accessed through a whole-word route and those accessed through a decomposable route. The former suffixes are low in productivity but high in relative frequency, and tend to attach closer to a base word. The latter suffixes are quite the opposite: they are high in productivity but low in relative frequency, and tend to come after the first type of suffixes.

4.2 Ordering Paradoxes in Terms of Relative Frequency

In this section, we provide an account of the five cases of ordering paradox that we introduced earlier, *-mental, -mentation, -mentary*, *-mentative*, and *-ageous,* using the relative frequency theory.

Let us consider *-mental*, first. To calculate relative frequencies, we have to know the frequencies of derived and base words. There are 109 forms of *-mental* derivative in COCA. Among them, however, only eight forms have a base separable from *-mental*; for the rest of 101 words, we can hardly recognize the bases as in *elemental*, *experimental*, *detrimental*, *supplemental*, and *monumental*.[7] If a base of a word is not recognizable, we could treat the derived word as lexicalized, to which affixes could attach freely. For this reason, we exclude such inseparable cases of *-mental* derivatives from the discussion of ordering paradoxes.

On these eight "transparent" cases of *-mental* derivatives, we then calculated the frequencies of 'base+*ment*' relative to the frequencies of the base words. This is to see if the base+*ment* portion of the word is lexicalized or not. The values of relative frequency are presented below:

Table 3. Relative frequencies of [base+*ment*] in ¬*mental* derivatives

Base	F ①	Derived Word	F ②	Relative F ②/①	Derived Word
argue	18408	argument	22190	1.20	argumental
compart	5	compartment	2296	452.90	compartmental
depart	1485	department	72671	48.93	departmental
develop	33114	development	82825	2.50	developmental
environ	51	environment	42168	826.80	environmental
frag	35	fragment	1212	34.62	fragmental
govern	2789	government	172528	61.80	governmental
judge	45235	judgment	15116	0.33	judgmental

In Table 3, the relative frequency of *argument* (1.20) is calculated by dividing the frequency of *argument* (22190) by the frequency of *argue* (18480). The derived word *argument* is used more frequently than its base and thus the relative frequency of *argument* is greater than 1, Since the relative frequency is greater than 1, the word *argument* is accessed through a whole-word route. Now *argument* is processed like a lexical word, to which an affix is freely attached, even though it is a derived

[7] See Appendix 2 for the complete list of the words containing the offending five suffix combinations, *-mental, -mentation, -mentary, -mentative,* and *-ageous*.

word. Note that all the words derived with *-ment* except *judgment* have a relative frequency higher than 1[8], which means that these derived words are processed like a lexicalized word, and so a Level I affix can freely come after it.

Now let us consider the other examples containing *-ment*: *-mentation*, *-mentary*, and *-mentative*. We have found 49 *-mentation*, 41 *-mentary* and 8 *-mentative* derivatives in COCA. In most cases, however, it is hard to separate the suffixes from its base (see Appendix 2 for the list of the words), except *argumentation*, *fragmentation*, *compartmentation*, *fragmentary*, and *argumentative*. We saw already in Table 3 that the frequencies of *argument*, *fragment*, and *compartment*, from which these five words are derived, are higher than those of their base words, *argue*, *frag*, and *compart*. The relative frequencies of *argument*, *fragment*, and *compartment* are thus greater than 1, which makes these words processed like a lexicalized word accessed through a whole-word route.

Finally, let us examine *-ageous*. We saw earlier that this was the only form of ordering paradox that is associated with the suffix *-age*. Out of 13 *-ageous* derivatives we have found in COCA, only three are the cases where we can separate the base from the suffixes, *advantageous*, *rampageous*, and *umbrageous* (see Appendix 2 for the list). We calculated the relative frequencies of the base+age portion of the words to see if they processed like a lexicalized word. The calculated values of relative frequency are presented below:

Table 4. Relative frequency of [base+*age*] in *-ageous* derivatives

Base	**F ①**	**Derived Word**	**F ②**	**Relative F ②/①**	**Derived Word**
advance	15293	advantage	23061	1.50	advantageous
ramp	671	rampage	652	0.97	rampageous
umbra	135	umbrage	162	1.2	umbrageous

Table 4 shows that all *-age* suffixed words have the relative frequencies higher than or very close to 1. Although the relative frequency of the word *rampage* is less than 1, it is still very close to 1. Since the values of relative frequency reflect the

[8] The relative frequency of *judgment* is low presumably because the base portion of the word, *judge*, is also frequently used as a noun meaning 'one who judges, especially in a court.' Note that the base of *-ment* suffixation must be a verb, rather than a noun.

cognitive process involved in perceiving morphologically complex words, we could say that the word *rampage*, which has 0.97 value of relative frequency, is considered to be acceptable as a lexicalized word.

In this section, we have found that there are certain Level II suffixes that come before Level I suffixes: *-age* and *-ment*. Relative frequency has been presented to provide an explanation of those ordering paradoxes. We have seen that base+*ment* and base+*age* forms are used more frequently than the base words, and thus their relative frequencies are higher than 1. The base+*ment* and base+*age* forms are therefore accessed through a whole-word route and should be considered lexical words rather than derived words. This then makes it possible for other suffixes to be attached to such words regardless of their level.

5. Conclusion

Based on the notion of relative frequency, this paper has provided a way to deal with ordering paradoxes that occur in English suffixation. We have first chosen 26 suffixes and classified them into Levels based on whether they undergo phonological changes such as stress shift and segmental changes (like Trisyllabic Shortening, Spirantization and Velar Softening). We have then examined how those suffixes are actually combined in terms of Levels based on the 110 attested suffix combinations taken from the CELEX lexical database. Although most of the suffixes follows the well-formed order of combinations as posited by LOH, some suffix combinations do not because here Level II suffixes precede Level I suffixes. We have found that the suffixes *-ment* and *-age* are always involved in these cases of ordering paradox.

This paper has then sought an explanation of ordering paradoxes based on the concept of productivity. We have seen, however, that productivity alone cannot be a solution for ordering paradoxes, much less the attested suffix orders.

We have thus investigated an account of ordering paradoxes from a different perspective which draws on the notion of relative frequency proposed by Hay (2002, 2003). When the relative frequency of a derived word is greater than 1, the word is accessed through a whole-word route, and the word is not considered to be a derived word, but rather a lexicalized word. It turned out that ordering paradox is a phenomenon where the base plus the first suffix of the offending suffix combinations

(i.e. *-ment* or *-age*) has its relative frequency higher than 1, and is thus processed like a lexicalized word so that any suffix, even Level I, can be attached to this portion of the word.

Appendix 1
Attested Suffix Combinations
(Obtained from CELEX)

(A) Base + I* (55)

Suffix Combinations				Example words
-al I	-ian I			episcopalian
	-ism I			naturalism
	-ist I			accidentalist
	-ity I			nationality
	-ize I			naturalize
	-ist I	-ic I		naturalistic
	-ize I	-ate I	-ion I	naturalization
-an I	-ism I			republicanism
	-ity I			Christianity
-ary I	-ize I			notarize
-ate I	-ion I			initiation
	-ive I			initiative
	-ion I	-al I		gravitational
	-ion I	-ism I		collaborationism
	-ion I	-ist I		collaborationist
-ic I	-al I			dialectical
	-an I			dialectician
	-ate I			metricate
	-ism I			romanticism
	-ist I			hypnotist
	-ize I			romanticize
	-ate I	-ion I		metrication
-ify I	-ate I			certificate

	-ic I			horrific
	-ion I			certification
-ion I	-al I			institutional
	-ary I			deflationary
	-ate I			affectionate
	-ic I			thermionic
	-ism I			expressionism
	-ist I			expressionist
	-al I	-ism I		Congregationalism
	-al I	-ist I		educationalist
	-al I	-ity I		conventionality
	-al I	-ize I		institutionalize
	-ist I	-ic I		exhibitionistic
	-al I	-ize I	-ion I	institutionalization
-ist I	-ic I			artistic
	-ic I	-al I		humoristical
	-ic I	-ate I		sophisticate
-ity I	-an I			puritan
	-ary I			hereditary
	-ive I			authoritatively
	-an I	-ical I		puritanical
	-an I	-ism I		puritanism
-ive I	-al I			substantival
	-ism I			collectivism
	-ist I			prescriptivist
	-ity I			objectivity
	-ize I			passivize
	-ate I	-ion I		captivation
	-ize I	-ate I	-ion I	collectivization
-ize I	-ate I	-ion I		capitalization
-ory I	-al I			purgatorial
-ous I	-ity I			nebulosity

(B) Base + II* (13)

Suffix combinations				Example words
-able II	-ly II			commendably
	-ness II			respectableness
-age II	-able II			marriageable
-ful II	-ly II			wonderfully
	-ness II			boastfulness
-ish II	-ly II			bearishly
	-ness II			foolishness
-less II	-ly II			aimlessly
	-ness II			carelessness
-ly II	-hood II			likelihood
	-ness II			timeliness
-some II	-ly II			handsomely
	-ness II			lightsomeness

(C) Base + I* + II* (33)

Suffix combinations					Example word
-al I	-ly II				naturally
	-ness II				naturalness
	-ist I	-ic I	-al I	-ly II	nationalistically
-an I	-ship II				custodianship
-ary I	-ly II				arbitrarily
	-ness II				arbitrariness
-ate I	-ly II				considerately
	-ness II				considerateness
-ic I	-al I	-ly II			academically
	-al I	-ness II			symmetricalness
-ic I	-al I	-ly II			parabolically
-ify I	-able II				classifiable
	-able II	-ly II			justifiably
	-ic I	-al I	-ly II		terrifically

-ion I	-able II				objectionable
	-less II				emotionless
	-ship II				relationship
	-al I	-ly II			institutionally
	-ate I	-ly II			affectionately
	-less II	-ly II			emotionlessly
	-less II	-ness II			emotionlessness
	-ist I	-ic I	-al I	-ly II	impressionistically
-ist I	-ic I	-ally II			realistically
-ity I	-ary I	-ly II			hereditarily
	-ive I	-ly II			authoritatively
	-an I	-ic I	-al I	-ly II	puritanically
-ive I	-ly II				cumulatively
	-ness II				massiveness
-ize I	-able II				realizable
	-ment II				aggrandizement
-ory I	-ly II				satisfactorily
-ous I	-ly II				ruinously
	-ness II				courageousness

(D) Base + II + I* (6)

Suffix combinations				Example word
-age II	-ous I			rampageous
-ment II	-al I			environmental
	-ary I			complementary
	-al I	-ist I		environmentalist
	-ate I	-ion I		argumentation
	-ate I	-ive I		argumentative

(E) Base + II + I* + II (3)

Suffix combinations				Example word
-age II	-ous I	-ly II		advantageously
-ment II	-al I	-ly II		environmentally
	-ate I	-ive I	-ly II	argumentatively

Appendix 2

Examples of Ordering Paradoxes (Including Apparent Cases; Obtained from COCA)

(A) *-mental* Derivatives

agroenvironmental	antienvironmental	antigovernmental	antisacramental
antisentimental	argumental	augmental	bioenvironmental
blumental	compartmental	decremental	departmental
detrimental	developmental	documental	ecodevelopmental
elemental	emmental	environmental	excremental
experimental	extramental	firmamental	fragmental
fromental	governmental	hyomental	hyperjudgmental
hypersentimental	impedimental	implemental	inclemental
incremental	instrumental	integumental	interdepartmental
interdevelopmental	interenvironmental	intergovernmental	intermental
intersegmental	intertestamental	intracompartmental	intradepartmental
intragovernmental	intramental	intrumental	judgemental
judgmental	microenvironmental	monumental	multicompartmental
multidepartmental	multisegmental	neogovernmental	neurodevelopmental
nidamental	nondepartmental	nondevelopmental	nonenvironmental
nonexperimental	nongovernmental	nonincremental	noninstrumental
nonjudgemental	nonjudgmental	nonornamental	nonsacramental
nonsentimental	nonsupplemental	nutrimental	ornamental
palaeoenvironmental	peramental	pimental	postexperimental
prodevelopmental	proenvironmental	regimental	regimental
rudimental	sacramental	scentimental	sedimental
segmental	semiexperimental	semigovernmental	sentimental
simmental	sociodevelopmental	socioenvironmental	sternomental
submental	subsegmental	supersegmental	supersentimental
supplemental	suprasegmental	tegmental	temperamental

undetrimental	unexperimental	unicompartmental	unjudgmental
unmonumental	unsentimental	vestimental	vestmental
viromental			

(B) *-mentation* Derivatives

alimentation	argumentation	augmentation	bioaugmentation
cementation	cofermentation	coimplementation	compartmentation
complementation	defragmentation	depigmentation	documentation
experimentation	fermentation	filamentation	flagmentation
fomentation	fragmentation	hyperalimentation	hyperpigmentation
hypopigmentation	implementation	implimentation	incrementation
instnumentation	instrumentation	intrumentation	istrumentation
lamentation	mentation	microinstrumentation	micropigmentation
nonimplementation	ornamentation	oversedimentation	oversegmentation
pigmentation	postimplementation	preimplementation	ragmentation
regimentation	reimplementation	resegmentation	sedimentation
segmentation	subfragmentation	supplementation	undersegmentation
viscosupplementation			

(C) *-mentary* Derivatives

alimentary	antiparliamentary	comentary	commentary
complementary	complimentary	documentary	elementary
extraparliamentary	figmentary	filamentary	fragmentary
implementary	interparliamentary	intertestamentary	metacommentary
metasedimentary	minidocumentary	momentary	myofilamentary
nonalimentary	noncomplimentary	nondocumentary	nonelementary
nonfragmentary	nonsedimentary	parliamentary	pigmentary
pseudodocumentary	rockumentary	rudimentary	sacramentary
sedimentary	segmentary	semidocumentary	supplementary
testamentary	uncomplementary	uncomplimentary	unparliamentary
vestimentary			

(D) *-mentative* Derivatives

argumentative	commentative	nonargumentative	overargumentative
augmentative	fermentative	nonfermentative	unargumentative

(E) *-ageous* Derivatives

advantageous	artrageous	contageous	courageous
disadvantageous	noncourageous	nutrageous	outrageous
pantageous	rageous	rampageous	umbrageous
uncourageous			

References

Allen, Margaret R. 1978. *Morphological investigations*. Dotoral dissertation, University of Connecticut.

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.

Aronoff, Mark. & Sridhar, Shikaripur N. 1983. "Morphological levels in English and Kannada, or Atarizing Reagan." In: "Papers from the Parasession on the Interplay of Phonology, Morphology and Syntax." *Chicago Linguistic Society* 19, 3-16.

Baayen, R. Harald. 1989. *A corpus-based approach to morphological productivity: Statistical analysis and psycholinguistic interpretation*. Doctoral dissertation, University of Amsterdam.

Baayen, R. Harald., Piepenbrock, Richard. & Gulikers, Leon. 1995. *The CELEX lexical data base(CDROM)*. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Chomsky, Noam, & Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Davies, Mark. 2008. Corpus of Contemporary American English. Available at: http://corpus.byu.edu.

Fabb, Nigel. 1988. "English suffixation is constrained only by selectional restrictions." *Natural Language and Linguistic Theory* 6, 527-539.

Giegerich, Heinz J. 1999. *Lexical strata in English: Morphological causes, phonological effects*. Cambridge: Cambridge University Press.

Hay, Jennifer B. 2002. "From speech perception to morphology: affix ordering revisited." *Language* 78:2, 527-555.

Hay, Jennifer B. 2003. *Causes and consequences of word structure*. London: Routledge.

Hay, Jennifer B. & Baayen, R. 2002. "Parsing and productivity." In Booij, Geert, & van Marle, Jaap Eds. *Yearbook of morphology 2001*, 203-235. Dordrecht: Foris Publications..

Hay, Jennifer B. & Plag, Ingo. 2004. "What constrains possible suffix combinations?: On the Interaction of grammatical and processing restrictions in derivational morphology." *Natural Language and Linguistics Theory* 22, 565-596.

Kiparsky, Paul. 1982. "Lexical phonology and morphology." In Linguistic Society of Korea Eds. *Linguistics in the morning calm*. 3-91. Seoul: Hanshin.

Marchand, Hans. 1969. *The categories and types of present-day English word formation: A synchronic-diachronic approach*. Munich: C. H. Beck.

Mohanan, Karuvannur M. 1986. *The theory of lexical phonology*. Dordrecht: Reidel.

Plag, Ingo. 2002. "The role of selectional restrictions, phonotactics and parsing in constraining suffix ordering in English." In Booij, Geert, and van Marle, Jaap Eds. *Yearbook of Morphology* 2001. 285-314.

Selkirk, Elisabeth. 1982. *The phonology of words*. Cambridge, MA: MIT press.

Siegel, Dorothy. 1974. *Topics in English morphology*. Doctoral dissertation, MIT. (Published 1979 by Garland: New York).

Spencer, Andrew. 1991. *Morphological theory*. Oxford: Blackwell.

Strauss, Steven L. 1982. "On 'relatedness paradoxes' and related paradoxes." *Linguistic Inquiry* 13, 694-700.

Szpyra, Joranta. 1989. *The phonology-morphology interface: cycles, levels and words*. London: Routledge.

Ju-Eun Lee · Sung-Hoon Hong
Department of English Linguistics
Hankuk University of Foreign Studies
270 Imun-dong, Dongdaemun-gu
Seoul 131-792
E-mail: hongshoon@hufs.ac.kr

Received: 2011. 03. 11
Revised: 2011. 04. 05
Accepted: 2011. 04. 22