WEIGHT AS A LINGUISTIC VARIABLE*
–WITH REFERENCE TO ENGLISH–

Nahm-Sheik Park

The present study claims that weight is a major linguistic variable crucial to the explanation of such linguistic phenomena as stress, length, order, and rhythm. It also suggests that the weight variable plays a key role in language processing. Although our discussion is confined mostly to English, most of our claims may be applicable mutatis mutandis to other languages as well.

0. Introduction

We will show in this paper that weight is a major linguistic variable. We will specifically demonstrate that it is of direct relevance to the assignment of stress as well as to the length and distribution of linguistic elements. Although almost all of my evidence is from English, most of my claims here are arguably applicable mutatis mutandis to other languages as well.

The variable of weight, as used in this paper, is definable in intuitive semantic terms in most instances. X may be said to be semantically heavier than Y if X is more substantive in meaning than Y in one way or another. For example, *Korea* is heavier than *Seoul* in that the former refers to a larger real-world entity than the latter. *Woman* is heavier than *person* in that the former takes more semantic features to define than the latter. *Dark* in *darkroom* is heavier than *room* in the same word in that *dark* figures more prominently in the meaning of the whole compound than *room* does.

Weight is also definable in structural terms. X may be said to be heavier than Y if X is longer or more complex structurally than Y. For example, *for that reason* is heavier than *therefore*, which in turn is heavier than *so*. For *for that reason* is longer and more complex than *therefore*, which in turn is

*The present study was supported by a research grant (1988-1989) from the Ministry of Education. An earlier version of this paper was read at the Spring/1981 Meeting of the Linguistic Society of Korea. I am grateful to Choon-Hak Cho, Dae-Sik Min, and Julian Hadland for their comments on the pre-publication draft of this paper. I am also grateful to my informants Carol Kreidler, Charles W. Kreidler, Duane L. Vorhees, and Joe Wood.

longer and more complex than *so*.

Our discussion of the weight variable will proceed in four main stages. Chapter 1 deals with weight as a variable in stress assignment. Chapter 2 treats the role of weight in determining the length of linguistic elements. Chapter 3 discusses weight as a determinant of word order. Chapter 4 addresses the role of weight as a determinant of rhythm.

1. Weight and Stress

1.0. The Weight-Over-Stress Principle

This principle, which governs stress assignment, may be stated as follows.

> THE HEAVIER A LINGUISTIC ELEMENT IS, THE MORE STRONGLY STRESSED IT IS; CONVERSELY, THE LIGHTER A LINGUISTIC ELEMENT IS, THE MORE WEAKLY STRESSED IT IS.

This principle guarantees that elements of greater communicative value are more strongly stressed and thus more distinctly pronounced than those of lesser communicative value. Thus it helps ensure an optimal transfer of information from speaker to hearer.

1.1. Content vs. Function Words

The weight-over-stress principle helps explain a wide range of stress phenomena. One of them is that content words tend to be more prominently stressed than function words. Note that this is in line with the weight-over-stress principle. For content words are generally heavier of semantic content than function words are.

However, function words are not always weakly stressed, and neither are content words always strongly stressed. We will show below that in neither case is the weight-over-stress principle really violated. Let us begin our discussion here with the following example.

(1) A: John is tired.[1]
B: So'm *Í*.

[1]The acute, cricumflex, grave, and breve marks are used in this paper to indicate primary, secondary, tertiary, and zero stress respectively.

I is a function word and thus light of inherent semantic content so that it normally gets weak stress. However, it gets strong stress in the above exchange in apparent violation of the weight-over-stress principle. The key to the puzzle here is to be found in the fact that *I* is contrasted with *John* so that it gains extra semantic weight from this contrastive emphasis. This means that *I* here is semantically heavy and thus qualifies for strong stress in line with the weight-over-stress principle.

Let us consider the following exchange, in which the function word *am* gets prominent stress in apparent contradiction of the weight-over-stress principle.

(2) A: You look hungry.
B: I *ám* hungry.

The function word *am* is light of inherent semantic content and thus normally gets weak stress. However, it gets affirmative emphasis in the above exchange so that it gains considerable semantic weight. Thus the strong stress on *am* in the above exchange is in line with the weight-over-stress principle, not in violation thereof.

Note at this point that clause-final function words such as *am* in the exchange below are always prominently stressed.

(3) A: Are you Korean?
B: Yes, I *ám*.

The clause-final *am* here is short for the longer phrase *am Korean* so that it carries not just its own weight but also the weight of *Korean*. Thus *am* here is quite heavy and qualifies for prominent stress in accordance with the weight-over-stress principle. Phenomena of this kind are discussed in great detail in Park (in preparation).

We are now in a position to explain in a principled way why depreposition al particles are normally more prominently stressed than their prepositional sources. Let us consider the two tokens of *in* in the exchange below.

(4) A: Is Bob *ĭn* the office?
B: Yes, he's *ín*.

The first token of *in* here is light of semantic content while the second is heavy thereof. Note that the first token is a mere preposition, which is a

function word, while the second token is a particle short for the longer phrase *in the house*. The particle *in* here thus carries not just its own weight but also the weight of *the house*. As a result, *in* here is normally heavier and thus more prominently stressed as particle than as preposition, which is in line with the weight-over-stress principle.

We can similarly explain why possessive pronouns tend to be more prominently stressed than their possessive-adjectival counterparts. Let us consider the two tokens of *his* in the paraphrase pair below.[2]

(5) a. That's *hĭs* book.
b. That's *hís*.

The second token of *his* here is short for the longer phrase *his book* so that it is normally heavier than the first token of the same. Thus the second token of *his* here normally gets stronger stress than the first in compliance with the weight-over-stress principle.

Conjunctions are function words and thus normally get weak stress. They may get strong stress, however, when they are short for a longer phrase and thus heavy. The second token of *if* in the exchange below is a case in point.

(6) A: He'll do it *ĭf* you pay him a lot.
B: Ah, yes. *Íf*.

Needless to say, the strong stress on the second token of *if* here is a consequence of the weight-over-stress principle.

Note in this connection that examples such as the following attest to the psychological reality of trace.

(7) A: Did you talk *tŏ* anybody?
B: No, I had nobody to talk *tò*.

The clause-final *to* in B's response here may be thought of as being followed by the trace of *nobody*. Thus this *to* is arguably short for the longer phrase consisting of itself plus the trace of *nobody* so that it is normally heavier than the first token of *to* here. Along this line of reasoning, we may contend that the second token of *to* here normally gets stronger stress than the first in keeping with the weight-over-stress principle.

[2]We will say that two expressions are paraphrases of each other if they can be transformationally related and thus assigned a cognitively similar semantic interpretation.

Function words may often receive prominent stress when they double as content words. Let us take for example *it*, as used in the following sentences.

(8) a. You're *ít*. (=You're *the next person (to play)*.)
b. She's got *ít*. (=She's got *sex appeal*.)

Normally *it* is a function word and thus light of semantic content so that it is weakly stressed. In the two sentences above, however, it is a content word, as is evident from the parenthesized glosses, and thus heavy of semantic content. Thus the prominent stress on *it* here is in line with the weight-over-stress principle, not in violation thereof.

Reflexive pronouns can also double as content words, as can be seen from (9b) below.

(9) a. He felt *hǐmsèlf*. (*himself*=a purely anaphoric object of *felt*)
b. He felt *hǐmsélf*. (*himself*="normal/well")

The reflexive *himself* is purely functional in (9a) while it has the contentive meaning of "normal" or "well" in (9b). Thus it is light of inherent semantic content in (9a) while it is heavy thereof in (9b). As a result, *himself* is normally (far) more prominently stressed in (9b) than in (9a) in accordance with the weight-over-stress principle.

Conjunctions may also double as content words, as in (10) below, so that they may get prominently stressed in accordance with the weight-over-stress principle.

(10) His proposal is full of *ífs* and *búts*.

It is interesting that one and the same function word *by* is more prominently stressed in (11a) below than in (11b).

(11) a. She wants a book *bỳ* herself. (*by*="near")
b. She wants a book *bý* herself. (*by*="written by")

Note that *by* is spatial in (11a) while it is non-spatial in (11b). Thus it is referentially more concrete in (11a) than in (11b) so that it is arguably heavier in the former than in the latter. If this is correct, then we can say that *by* is more prominently stressed in (11a) than in (11b) in accordance

with the weight-over-stress principle.

Incidentally, weight due to referential concreteness often translates not into stress but into length. For example, *by the car* and *on the television* are referentially more concrete than *by ϕ car* and *on ϕ television* respectively so that the former are heavier and thus longer than the latter. This is in line with the weight-over-length principle to be discussed in Chapter 2.

Content words such as *person* in (12b) below normally receive relatively weak stress in apparent violation of the weight-over-stress principle.

(12) a. He's a *brîlliănt dóctŏr.*
b. He's a *brílliănt pèrsŏn.*

Nuclear stress normally falls on the tonic syllable of the last content word in a tone unit. Thus it normally falls on the tonic syllable of *doctor* in (12a). In (12b), however, it does not normally fall on the tonic syllable of *person*, which is the last content word in the sentence. This is due to the inherent semantic levity of *person*, which is almost as light as a personal pronoun. Thus it is not really in violation of the weight-over-stress principle that *person* does not normally get nuclear stress in (12b).

Incidentally, *person* is not the only content word that is light of inherent semantic content. All other general-purpose content words are semantically light and thus normally get relatively weak stress. Among these are nouns such as *thing, matter, stuff,* and *guy* as well as verbs such as *come, go, get, grow, put, set, lie, sit, stand, say,* and *make.*

Content words also get relatively weak stress when they are collocationally predictable and thus light of information content. Let us compare the two sentences below.

(13) a. The *béll* is *ringĭng.*
b. The *bêll* is *glíttĕrĭng.*

Note that ringing is a bell's reason for being so that mere mention of a bell automatically reminds us of its ringing function. Thus *bell* renders *ringing* collocationally predictable when they are in a subject-verb relationship, as in (13a). As a result, *ringing* in (13a) is so light of information content that it does not normally qualify for nuclear stress.

On the other hand, there is no such close link of association between a bell and glittering so that *bell* does not render *glittering* collocationally predictable in (13b). Thus *glittering* in (13b) retains its original contentive

weight and thus normally qualifies for nuclear stress. Note that the normal de-stressing of *ringing* here, as opposed to its disallowance in *glittering*, is a consequence of the weight-over-stress principle.

Content words also tend to get de-stressed when they refer to the immediate context of speech. Among such content words are *here* and *now*, as used in the following sentences.

(14) a. I work *hère.*
b. I'm okay *nòw.*

Note that *here* and *now* refer to the immediate context of speech so that their referents are readily accessible to the interlocutors' consciousness. For this reason, *here* and *now* are usually light of information content, that is, unless they gain additional semantic weight, say, from contrastive emphasis. Thus *here* and *now* normally get weak stress in compliance with the weight-over-stress principle.

One and the same content word may vary in stress from one sense to another, as can be seen from the two sentence pairs below.

(15) a. He bought *a fêw* books.
b. He bought *féw* books.
(16) a. She ate *a lîttĭle* rice.
b. She ate *líttĭle* rice.

A few and *a little* are positive while *few* and *little* are negative so that the former are semantically lighter than the latter. As a result, *few* is normally less prominently stressed in (15a) than in (15b) while *little* is less prominently stressed in (16a) than in (16b). Note that this is all in keeping with the weight-over-stress principle.

Let us now turn our attention to a large class of constructions, which I will call governed constructions for want of a better name. A governed construction comprises two immediate constituents, i.e. a governor followed by a governee, as in the following examples.

(17) a. preposition + complement: *Tŏ peáce!*
b. subordinator + complement: *Ĭf póssĭble.*
c. auxiliary + main verb: *căn dó*
d. verb + object: *Câll Jóhn.*
e. determiner + noun: *thĕ Lórd*
f. title + name: *Mîss Smíth*

g. adjective + noun: *goôd schoôls*
h. intensifier + intensified: *vêrỳ wéll*

Note that governors tend to be on the functional side while governees tend to be on the contentive side. Thus governors tend to be semantically lighter than governees so that the former normally receive weaker stress than the latter. Note that this is in compliance with the weight-over-stress principle.

1.2. Compound Words

Conpound words also obey the weight-over-stress principle. Let us take the following compound nouns for example.

(18) a. *gréen*hòuse
b. *fíre*flỳ

In either compound here, the first constituent word figures more porminently in the meaning of the whole compound than does the second. Thus the first constituent word is semantically heavier than the second so that the former is more prominently stressed than the latter in accordance with the weight-over-stress principle.

Let us consider the following compound nouns.

(19) a. flát*wàre*, sílvěr*wàre*, eárthěn*wàre*
b. blúe*fish*, góld*fish*, swórd*fish*

The compounds in either set here share the same second consituent while they have different first constituents. Thus they are contrasted with each other with respect to the first constituents so that these first constituents are semantically heavier than the shared second constituent. In fact, the first cosntituent in each compound here is like a stem while the second constituent is like a suffix. Thus the first constituent receives stronger stress than the second in accordance with the weight-over-stress principle.

Note at this point that not all compound nouns get primary stress on the first constituent. Let us compare the two compounds in either pair below.

(20) a. lífe stòrỳ
b. lîfe ĭmprísŏnmĕnt

(21) a. Sússéx màn
b. Sûsséx víllăge

In (20) above, *life* is semantically heavier and thus more strongly stressed than *story* while it is semantically lighter and thus less strongly stressed than *imprisonment*. In (21), *Sussex* is semantically heavier and thus more strongly stressed than *man* while it is semantically lighter and thus less strongly stressed than *village*. Thus, whichever constituent is more strongly stressed, all the compounds here follow the weight-over-stress principle.

The difference in stress pattern between numerals in *-teen* and those in *-ty* can also be explained along similar lines. Let us take the two numerals below for example.

(22) a. fíftéen (=five *and* ten)
b. fíftŷ (=five *times* ten)

Note that *fif-* and *-teen* are coordinate and thus coequal terms in *fifteen* while *fif-* is superordinate to *-ty* in *fifty* in that the former operates on the latter. Thus *fif-* and *-teen* are about the same weight in *fifteen* while *fif-* is heavier than *-ty* in *fifty*. As a result, the stress is about equally strong on *fif-* and *-teen* in *fifteen* while it is stronger on *fif-* than on *-ty* in *fifty*. Note that this is all in line with the weight-over-stress principle.

Note here that *fifteen*, pronounced in isolation, gets more prominent stress on *-teen* than on *fif-*. Since *-teen* is of greater numerical value than *fif-*, this is evidently a consequence of the weight-over-stress principle. Note also that *fifteen* may get stronger stress on *fif-* than on *-teen* when it is contrasted with such numerals as *sixteen*, as in serial counting. This is becuase *fif-* gets extra semantic weight from the contrastive emphasis here. This is also a consequence of the weight-over-stress principle. We will come back to the contrast between *-teen* and *-ty* numerals in 4.1., where we will discuss why *-ty* is always atonic while *-teen* is always tonic.

1.3. Stems and Affixes

Stems are normally heavier of semantic content than affixes so that the former customarily get stronger stress than the latter in accordance with the weight-over-stress principle. This can be seen from examples such as the following.

(23) a. ùn*líke*lȳ
b. prò-*líf*ĕr
c. *lóne*lĭnĕss

Of affixes, prefixes tend to get stronger stress than suffixes, as can be seen from (23a) and (23b). The prefixes *un-* and *pro-* here are semantically heavier and thus more prominently stressed than the suffixes *-ly* and *-er* respectively. This is in line with the weight-over-stress principle.

The weight-over-stress principle also helps explain why superficially identical prefixes are often differentially stressed, as in the following example.

(24) rè-sígn vs. rĕsígn

The prefix *re-* is semantically heavier in *re-sign* than in *resign* so that it gets stronger stress in the former than in the latter, which is in line with the weight-over-stress principle.

Prefixes may even receive primary stress under contrastive emphasis. For example, *offense* and *defense* contrast with each other in the context of, say, American football so that *of-* and *de-* get contrastive emphasis. Thus the prefixes here gain exra informational weight and thus may get primary stress in accordance with the weight-over-stress principle.

Note at this point that suffixes are not always zero-stressed, as can be seen from the following examples.

(25) a. Chin*ése*
b. engin*éer*
c. less*ée*
d. Paul*ìne*
e. social*ìte*
f. social*ìze*
g. beauti*fȳ*
h. Kennedy*ésque*
i. lemon*áde*
j. verb*óse*

The suffixes in (25) are all phonologically heavy in that they contain either a complex vowel or a consonant cluster. Many of them are also etymologically heavy in that they orignate in heavy French suffixes. Thus phonological (and etymological) weight is apparently responsible for the exceptionally prominent stress that the suffixes here receive. Note that this prominent suffixal

stress is a consequence of the weight-over-stress principle.

The suffix *-ose* in (25j), for one, may be semantically heavy as well in that it has the contentive meaning of "full of." This semantic weight may also be responsible for the prominent stress that the suffix receives.

It is interesting that *-ine* in (25d) gets stronger stress as a feminine suffix than as an adjectival suffix. Note that the feminine gender is marked while the masculine gender is unmarked so that femininity is heavier of inherent semantic content than masculinity. Besides, the feminine name *Pauline* contrasts with the masculine name *Paul* so that the feminine suffix *-ine* receives constrastive emphasis. Thus *-ine* is semantically heavier and thus more prominently stressed as a feminine suffix than as an adjectival suffix, which is in line with the weight-over-stress principle.

Note in this connection that the strong stress on *-ette* in (26) below may result at least in part from its relatively heavy semantic weight.

(26) a. usher*étte*
b. room*étte*
c. leather*étte*

The suffix here not only originates in a heavy French suffix but also has the contentive meaning of "small," "female," or "fake." Thus it is heavy etymologically as well as semantically, so that the prominent stress it receives is in line with the weight-over-stress principle.

Suffixes may also get prominent stress when they receive contrastive emphasis. The suffix *-ess* in the sentence below is a case in point.

(27) I met not just the *prince* but also the *princéss*.

Note that *princess* is contrasted with *prince* here so that contrastive emphasis falls on *-ess*. Thus *-ess* here is heavy so that it qualifies for prominent stress in accordance with the weight-over-stress principle.

Note here that the suffix *-ess* has the contentive meaning of "female" so that it is fairly heavy of semantic content as suffixes go. It also originates in a heavy French suffix. Note also that *princess* customarily contrasts with *prince* in monarchies such as Britain. It is evidently for these reasons that British English assigns primary stress to the second syllable of *princess*. Note that this is also in line with the weight-over-stress principle.

The optional primary stress on the suffix *-or* in the examples below is also attributable to contrastive emphasis.

(28) a. less*ŏr*
b. less*ór* (as opposed to *lessee*)
(29) a. vend*ŏr*
b. vend*ór* (as opposed to *vendee*)

Lessor and *vendor* are often contrasted with *lessee* and *vendee* respectively, especially in legalese, so that *-or* often gets extra semantic weight from the contrastive emphasis. This extra semantic weight makes the suffix eligible for the prominent stress that it often receives. Note that this optional primary stress on *-or* is in line with the weight-over-stress principle.

Note at this point that derivational suffixes are usually less light of semantic content and thus less weakly stressed than inflectional suffixes. Note further that even inflectional suffixes may sometimes get some extra semantic weight and thus receive less weak stress than is normally the case. Let us take the singular-plural pairs below for example.

(30) a. cris*is* vs. cris*es*
b. oas*is* vs. oas*es*

The plural ending *-es* gets contrastive emphasis here in that it contrasts with the singular ending *-is* so that *-es* is heavier than *-is*. Further, *-es* is also heavier than *-is* in that plurality is inherently heavier than singularity. Thus the heavier *-es* here gets less weak stress than the less heavy *-is* in accordance with the weight-over-stress principle.

2. Weight and Length

2.0. The Weight-Over-Length Principle

There is a high degree of correlation between the weight of a linguistic element and its length, which is summed up in the weight-over-length principle given below.

> THE HEAVIER A LINGUISTIC ELEMENT IS, THE LONGER IT TENDS TO BE; CONVERSELY, THE LIGHTER A LINGUISTIC ELEMENT IS, THE SHORTER IT TENDS TO BE.

Like the weight-over-stress principle, this principle makes sure that elements of greater communicative value are given more substantive expres-

sion than those of lesser communicative value. It is thus apparently instrumental in facilitating a maximally efficient transfer of information from encoder to decoder.

2.1. Content vs. Function Words

According to the weight-over-length principle, content words should normally be longer than function words. For the former tend to be heavier of semantic content than the latter. The principle seems to be borne out since the average content word is indeed longer than the average function word. For one thing, content words are often polysyllabic while function words rarely are. For another, even where both content and function words are monosyllabic, the former tend to be phonologically longer than the latter. Let us take the following pairs of monosyllabic words for example.

(31) a. four vs. for
b. two vs. to
c. can (=noun) vs. can (=auxiliary verb)

Note that in natural connected speech the vocalic nucleus is more distinct and thus longer in the first member of each pair here than in the second. This is because the first member is a content word while the second is a function word so that the former is semantically heavier than the latter.

The following pairs also show that monosyllabic words tend to be phonologically longer as content words than as function words.

(32) a. so, do (as in "*do,* re, mi...")
b. to, do (as in "I *do* not live here.")

Note that (32a) and (32b) comprise content and function words respectively. Note further that /ow/ is longer than /uw/, /u/, or /ə/ so that the words are longer in (32a) than in (32b). In our terms, this is because the content words of (32a) are semantically heavier than the function words of (32b).

Let us consider the following additional examples.

(33) a. (Santa) Fe, pe, (per) se, (in) re
b. he, she, we, me, be

The monosyllabic words here are contentive in (33a) and functional in (33b)

so that they are semantically heavier in the former than in the latter. This difference in semantic weight is apparently reflected in the vocalic nucleus, which ranges from /iy/ to /ey/ in (33a) and from /i/ to /iy/ in (33b). Since /ey/ is longer than /iy/, which in turn is longer than /i/, the vocalic nucleus tends to be longer in (33a) than in (33b). Thus the content words here also tend to be longer than the function words. Note that this, as well as what we have said regarding (31) and (32), accords with the weight-over-length principle.

2.2. Contractions

Function words often allow contraction to shorter forms while content words normally do not. Since function words are generally lighter of semantic content than content words are, this is in line with the weight-over-length principle. Let us consider the following paraphrase pairs.

(34) a. Who *does* that sort of thing?
b. *Who'*s* that sort of thing?
(35) a. What *does* that mean?
b. What'*s* that mean?

The contentive *does* is too heavy to allow contraction so that (34b) is ungrammatical, while the functional *does* is light enough to do so with the result that (35b) is perfectly grammatical.

Will is another word that normally allows contraction as a function word, but not as a content word. Let us consider the following paraphrase pairs.

(36) a. I must do whatever he *wills*.
b. *I must do whatever he'*lls*.
(37) a. I *will* go.
b. I'*ll* go.

Note incidentally that *will* is quite heavy semantically even as a function word when it denotes insistence. In fact, the *will* of insistence happens to be too heavy to allow contraction to -'*ll* so that *will* or its contraction -'*ll* in (37) cannot denote insistence. This again is a consequence of the weight-over-length principle.

Have is still another word that allows contraction as a function word, but not as a content word. The following paraphrase pairs illustrate this be-

havior of *have.*

(38) a. We *have* lunch at noon.
b. *We'*ve* lunch at noon.
(39) a. We *have* seen her already.
b. We'*ve* seen her already.

It is perhaps in order here to deal with certain constraints on the contraction of function words. Let us begin by considering the following paraphrase set.

(40) a. He *is not* tall.
b. He'*s not* tall.
c. He *isn't* tall.
d. *He'*sn't* tall.

Note that *isn't* may not shorten to *-'sn't* because it is already quite heavy in that it is a combination of two words, i.e. *is* and *not.* Note further that this is in line with the weight-over-length principle.

Note in this connection that the auxiliary *can* ordinarily contracts from /kæn/ to /k(ə)n/ because of its semantic levity. It cannot similarly contract in *can't,* however, because *can't* is a combination of two words, i.e. *can* and *not,* so that it is quite heavy. What we have said here about *can/can't* also applies to *are/aren't, have/haven't, has/hasn't, had/hadn't,* etc.

It is interesting to note here that British English regularly upgrades the vowel in *can* to /ɑ:/ when it combines with *-n't,* as shown below.

(41) a. I *can* go. (/kæn/ or /k(ə)n/)
b. I *can't* go. (/kɑ:nt/)

Note that /ɑ:/ is phonologically longer than either /æ/ or /(ə)/. Thus the nucleus difference between *can* and *can't* is more clearly reflective of their weight difference in British English than in American English. On this particular point, British English obeys the weight-over-length principle more faithfully than American English.

Note that *be* can normally contract in the present tense, but not in the past tense, as can be seen from the following paraphrase pairs.

(42) a. He *is* here.
b. He'*s* here.

(43) a. He *was* here.
b. *He*'s* here.

Was is much heavier than *is* in that it is a combination of *is* and the past tense. Thus we may say that *was* is too heavy to allow contraction while *is* is light enough to do so. Besides, the present-tense contraction -*'s* may block the hypothetical past-tense contraction -*'s*, not the other way around, because the former is more basic than the latter.

Another constraint on the contraction of function words has to do with function words under special emphasis. Function words may not contract under, say, affirmative emphasis. *Is* and *has* in the exchanges below are cases in point.

(44) A: He's not to blame, I tell you.
B: But he *is* to blame. (=*?But he*'s* to blame.)
(45) A: He hasn't even called.
B: But he *has* called. (=*?But he*'s* called.)

Is and *has* in B's responses here receive affirmative emphasis so that they are too heavy to allow contraction to -*'s*. This is, of course, a consequence of the weight-over-length principle.

The third and final constraint on the contraction of function words has to do with clause-final function words. Let us take for example the clause-final *is* and *will* in the parapharase pairs below.

(46) a. She isn't ready, but he *is*.
b. *She isn't ready, but he*'s*.
(47) a. He won't help us, but she *will*.
b. *?He won't help us, but she*'ll*.

Note that *is* and *will* here are short for, and thus carry the weight of, *is ready* and *will help us* respectively. Thus they are too heavy to allow contraction, which is why (46b) and (47b) are ungrammatical. This again is a consequence of the weight-over-length principle.

2.3. Long vs. Short Forms

The weight-over-length principle also helps explain why some expressions

have two allo-forms, one long and one short. Let us begin our discussion here by considering the following sentence pairs with respect to the words in italics.

(48) a. We saw him *there*. / ðɛər/
b. *There* was a girl near him. / ðər/
(49) a. Harvard is *the* university. / ðiy/
b. I like *the* university. / ðə/

The italicized word is semantically heavier in the first member of either pair here than in the second. For it has a contentive meaning in the first member while it has only a functional meaning in the second. Note that the italicized word is longer in the first member here than in the second in that the nucleus is more distinct and thus longer in the former than in the latter. This is, of course, a consequence of the weight-over-length principle.

The weight-over-length principle also makes available a rational explanation for the two allo-forms of the suffix *-ful*, as exemplified below.

(50) a. hand*ful*, mouth*ful*, cup*ful* /ful/
b. care*ful*, beauti*ful*, wonder*ful* /fəl/

Note that *-ful* is semantically heavier in (50a) than in (50b) in that the meaning of "full" is more salient in the former than in the latter. Note further that /u/ is more distinct and thus longer than /ə/ so that *-ful* is longer in (50a) than in (50b), which is in line with the weight-over-length principle.

A similar account is applicable to the two divergent forms of the etymologically identical suffix exemplified below.

(51) a. man*like*, god*like*, earth*like*
b. man*ly*, god*ly*, earth*ly*

The meaning of "like" is more salient in (51a) than in (51b) so that the suffix is semantically heavier in the former than in the latter. Note that the semantically heavier *-like* is phonologically longer than the semantically lighter *-ly*, which is in line with the weight-over-length principle.

The two alternative pronunciations of *-man*, as exemplified below, can also be explained along similar lines.

(52) a. ape*man*, super*man*, cave *man* /mæn/
b. post*man*, police*man*, line*man* /mən/

The meaning of "human being" is more prominent in (52a) than in (52b) so that *-man* is semantically heavier and thus phonologically longer in the former than in the latter. This, of course, accords with the weight-over-length principle.

The weight-over-length principle also offers insight into the alternation between *mouth* and *-mouth,* as shown below.

(53) a. *mouth* /m*aw*θ /
b. Ply*mouth,* Dart*mouth,* Ports*mouth* /məθ /

The meaning of "mouth" is far more salient in *mouth* than in *-mouth* so that the former is semantically much heavier than the latter. Note that the nucleus is longer in *mouth* than in *-mouth* so that the former is longer than the latter, which accords with the weight-over-length principle. It may be noted here that a similar explanation applies to the alternation between *land* and *-land,* as in *England* and *Scotland,* as well as that between *day* and *-day* /diy/, as in *Sunday* and *Monday.* It is interesting that *-day* in *weekday* is /dey/ because it contrasts with *-end* in *weekend* so that it receives extra semantic weight. Note incidentally that *-day* in *Monday,* etc., *-land* in *England,* etc., and *-mouth* in *Plymouth,* etc. are quasi-suffixal here. We will come back to this point in 2.4. in connection with (84) through (92).

The alternation between *-have* (as in *behave*) and *have* and between *parson* and *person,* as shown below, can also be explained in a similar manner.

(54) a. -have /h*ey*v/ vs. have /hæv/ or /həv/
b. parson /pár*s*ən/ vs. person /pə́rsən/

Note that *have* is light of semantic content, especially as an auxiliary, while *-have* has the contentive meaning of "control." Note further the /ey/ is longer than either /æ/ or /ə/ so that *-have* is longer than *have,* which accords with the weight-over-length principle. Note also that *parson* is heavier of semantic content than its cognate *person* and that the former is phonologically longer than the latter in that /ɑ/ is longer than /ə/. This also accords with the weight-over-length principle.

It is interesting that *farm* and *firm,* which are of the same ultimate etymological origin, differ in vocalic nucleus. The nucleus is more distinct and thus

longer in *farm* than in *firm* so that *farm* is phonologically longer than *firm*. Note here that in Middle English *farm* and *firm* were *ferme* and *ferm* respectively so that *farm* is of greater latent (phonological) weight than *firm*. We may thus argue that *farm* is longer than *firm* in accordance with the weight-over-length principle.

Note in passing here that the vowel laxes and shortens from /ey/ to /e/ in (55b) below, but not in (55a).

(55) a. p*ay* → p*ay*s/p*ai*d
b. s*ay* → s*ay*s/s*ai*d

The stem is semantically much lighter in (55b) than in (55a). Thus the semantically lighter stem here allows its vowel to shorten in the process of conjugation while the semantically heavier stem does not. Note that this is in accordance with the weight-over-length principle. Incidentally, the syllable closure is apparently also relevant to the shortening of the nucleus here. As we shall see in 4.3., the weight/length of a nucleus is in inverse ratio to that of the coda.

The weight-over-length principle also throws light on why the first word in each pair below is phonologically longer than the second.

(56) a. prayer /pr*eiə*r/
b. prayer /pr*ɛə*r/
(57) a. busyness /biz*i*nis/
b. business /biznis/
(58) a. anti-Arctic
b. Antarctic

Although both words in each pair here are etymologically polymorphemic, the second word is generally perceived as monomorphemic while the first is generally perceived as polymorphemic. Thus the first word in each pair here is arguably heavier than the second so that the former is longer than the latter in compliance with the weight-over-length principle. Note that we may say exactly the same thing about the pair *invalid* / invælid / and *invalid* /invəlid/.

It is interesting to note in this connection that *going to* may collapse into *gonna* in (59b) below, but not in (59a).

(59) a. I'm *going to* New York. → *I'm *gonna* New York.
b. I'm *going to* visit New York. → I'm *gonna* visit New York.

Note that (-'*m*) *going to* is semantically more than one unit in (59a) while it is arguably one unit in (59b) in that it is equivalent to *will*. Thus *going to* is much lighter in (59b) than it is in (59a) so that it may shorten and collapse into *gonna* in the former, but not in the latter. Note that this is in line with the weight-over-length principle.

Note in this connection that *want to* may collapse into *wanna* in "He's the one I *want to* talk to," but not in "He's the one I *want to* talk." Note that the logical subject of the infinitive is coreferential with that of *want* in the first sentence while it is not in the second. Thus *want to* is arguably lighter in the first sentence than in the second so that it may collapse into *wanna* in the former, but not in the latter, which is in line with the weight-over-length principle.

Let us now consider the two sentences below.

(60) a. I got good grades(,) and *I* made my parents happy.
b. I got good grades and ϕ made my parents happy.

The two sentences here arguably derive from a similar underlying structure, perhaps from something like the first sentence here. Note, however, that the two sentences here have different semantic interpretations in that the two events referred to are related to each other as cause and effect in the second sentence, but not in the first. Thus the first sentence contains two propositions, so to speak, while the second contains just one. As a result, the first sentence is semantically heavier than the second so that the former is longer than the latter in accordance with the weight-over-length principle.

We have seen in 1.1. that governors tend to be semantically lighter than governees. Note at this point that governors should be shorter than governees if the weight-over-length principle is to be upheld. In fact, this turns out to be the case in most instances, providing additional validation for the weight-over-length principle.

Note in this connection that one and the same word often has two alternative forms, one as governor and one as non-governor. In such cases, the governor often assumes a shorter form than the non-governor. Let us consider the following paraphrase pairs.

(61) a. *Mr.* Kim is my *master*.
b. **Master* Kim is my *Mr.*
(62) a. *Mrs.* Kim is my *mistress*.
b. **Mistress* Kim is my *Mrs.*

(63) a. *Mt.* Everest is a marvelous *mountain.*
b. **Mountain* Everest is a marvelous *Mt.*

Note that the governors here are semantically lighter than their respective non-governor forms. Thus we may say that *master, mistress,* and *mountain* shorten in governor position to *Mr., Mrs.,* and *Mt.* respectively under the pressure of the weight-over-length principle.

The following paraphrase pairs also show that governors tend to favor shorter forms.

(64) a. He owns a *gas* station.
b. (?)He owns a *gasoline* station.
(65) a. He majors in *U.S.* history.
b. (?)He majors in *United States* history.

Note that the governor slot in either pair here is filled with one of two allo-forms of the same word. Note further that the shorter of the two allo-forms is generally favored over the longer one here so that the first sentence in either pair above is more natural than the second. This preference for a shorter governor is, of course, in line with the weight-over-length principle. For governors are typically light of semantic content.

The following paraphrase pairs also show that governors tend to favor shorter forms.

(66) a. eggs that are *newly laid*
b. *new-laid* eggs
(67) a. a house with *three rooms*
b. a *three-room* house
(68) a. a scholar who was *educated in the United States*
b. a *U.S.-educated* scholar

Relevant to our discussion of the above data may be the fact that premodification is informationally lighter than postmodification (See 3.1.). Incidentally, our discussion here throws light on the derivation of the article *a(n)* from the numeral *one. One* may function as either governor or non-governor so that it is lighter in the former function than in the latter. It is in fact so light as governor that it has allowed shortening to the proclitic *a(n).* Note that this is a consequence of the weight-over-length principle. Note also that we can offer a similar account for the derivation of *the* and *through*

from *that* and *thorough* respectively. Perhaps similarly explainable is the derivation of the second alternant from the first in such Korean word pairs as /hana/~/han/ 'one', /tul/~/tu/ 'two', /ses/~/se/ 'three', /nes/~/ne/ 'four', and /sælop/~/sæ/ 'new'.

2.4. Deletion

The weight-over-length principle provides a rationale for the phenomenon of deletion, which is a form of shortening. Note that only elements of extreme levity may be deleted, as can be seen from the following paraphrase pairs.

(69) a. (?)She's far from *being* recovered.
b. She's far from ϕ recovered.
(70) a. (?)He's regarded as *being* overqualified.
b. He's regarded as ϕ overqualified.
(71) a. (?)It's close to *being* miraculous that he's still alive.
b. It's close to ϕ miraculous that he's still alive.

The two sentences in each pair here are such that the second sentence arguably derives from the first. Note here that the second sentence in each pair is slightly more natural than the first so that *being* is more often than not deleted from the first sentence. Note further that *being* may be deleted here because of its extreme semantic levity, which is in line with the weight-over-length principle.

It is interesting to note here that *be* is sometimes less likely to be deleted in its existential sense than in its copular sense. Let us consider the following paraphrase pairs.

(72) a. She seems *to be* asleep.
b. She seems ϕ asleep.
(73) a. She seems *to be* abroad.
b. *She seems ϕ abroad.

Note that *be* is a mere copula in (72a) while it has the fairly contentive meaning of "exist" in (73a) so that it is semantically heavier in the latter than in the former. Thus we can say that the copular *be* here may be deleted on account of its extreme semantic levity while the existential *be* may not because of its slightly greater weight. This is again in keeping with

the weight-over-length principle.

Note in this connection that *to* is deleted along with *be* in the derivation of (72b) from (72a) while it is not deleted along with *being* in the derivation of (71b) from (71a). This may be because the infinitive-marking *to* in (72a) is extremely light of semantic content while the directional *to* in (71a) is not quite as light thereof. If correct, this is also in line with the weight-over-length principle. Note incidentally that *to* is referentially more concrete and thus semantically heavier as a directional preposition than as an infinitive marker (See our comments on (11) in 1.1.).

It may be in order here to point out that the deletion of (*to*) *be* may sometimes entail a substantial change in meaning. For "I found him *to be* dead" is definitely different in meaning from "I found him ϕ dead."

It is like *be* in that it is also deletion-prone because of its semantic levity. Let us consider the following sentences, where *it* is optional.

(74) a. He made (*it*) clear that he would run for mayor.
b. He made (*it*) plain that he would run for mayor.
c. (*It*) being Christmas, the offices were closed.

It may be deleted here because of its semantic levity, which is in line with the weight-over-length principle.

Sentences such as the following may involve a similarly motivated *it*-deletion in their derivational history.

(75) a. As ϕ was usual, he woke up at five.
b. As ϕ so often happens, the machine is out of order again.
c. Spend this money as you see ϕ fit.

In each sentence here, we may quite plausibly posit an underlying *it* for the slot marked with ϕ. We may argue that this underlying *it* is so light of semantic content that it is deleted obligatorily in the process of deep-to-surface mapping here.

The idiomatic *as is*, as exemplified in (76) below, may also undergo a similar *it*-deletion in the course of its derivation.

(76) We bought the table *as* ϕ *is*.

We are suggesting here that (76) probably derives from "We bought the table *as it is*" and that *it* is deleted in the course of deep-to-surface mapping

on account of its semantic levity.

Note that expressions such as *as soon as possible* may involve the simultaneous deletion of *it* and *be*. For *as soon as possible*, for one, quite plausibly derives from *as soon as it is possible*. Needless to say, this deletion should also be a consequence of the weight-over-length principle.

The weight-over-length principle can also help explain the derivation of *please*, as exemplified in (77) below.

(77) Come this way, *please*.

The adverbial *please* here derives from *if you please*, which in turn derives from *if it please(s) you*. Note that the words deleted in the derivational process here are all light of semantic content so that the deletion here is arguably a consequence of the weight-over-length principle. In 3.6., we will come back to the derivation of *if you please* from *if it please(s) you* (See (225)).

The weight-over-length principle can also help explain the derivation of *how come*, as exemplified in the question below.

(78) *How come* you're so late?

This sentence apparently derives from something like "*How does it come (about) that* you're so late?" Assuming that this derivation is correct, we can see that the words deleted in the derivational process here are all light of inherent semantic content. Thus the deletion here is also apparently a consequence of the weight-over-length principle.

General-purpose words often undergo deletion on account of their inherent semantic levity, which is also in line with the weight-over-length principle. In support of this contention, we may cite the following sentences, in which the parenthesized words may be deleted.

(79) a. I must (*go*) away.
b. Truth will (*come*) out.
c. This is (*the place*) where I work.
d. That's (*the time*) when I called him.
e. What's (*the matter*) with you?[3]

[3]In 2.6., we will discuss another reason for the deletion of *the matter* in connection with (122b).

Let us turn our attention now to collocational lightening of information as a trigger for deletion. Let us begin our discussion here with the following sentences.

(80) a. Where did you *park (your car)?*
b. He *craned (his neck)* for a better view.

In either sentence here, the verb collocates with the object so customarily that, given the verb, the object is predictable. Thus the object here is almost empty of information content so that it may be deleted in accordance with the weight-over-length principle.

Collocational lightening of information is also responsible for the deletion illustrated by the first member of each sentence pair below.

(81) a. Jack *jumped (over)* the fence.
b. Jack *jumped under* the fence.
(82) a. Jack *climbed (up)* the hill.
b. Jack *climbed down* the hill.
(83) a. Jack *fled (from)* the country.
b. Jack *fled to* the country.

The collocational link between verb and preposition is so close in the first sentence of each pair here that, given the verb, the preposition is predictable. On the other hand, this is not the case in the second sentence of the same pair. Thus the preposition is light enough to be deleted in the first sentence of each pair here, but not in the second. Note that this is all in line with the weight-over-length principle.

Collocational lightening of information also triggers the deletion exemplified by the following sentences.

(84) a. The *butcher('s (shop))* is over there.
b. I met her at the *florist('s (shop)).*

The occupational nouns *butcher('s)* and *florist('s)* collocationally presuppose the locative noun *shop* here. Thus *shop* is extremely light of information content here so that it may be deleted. Incidentally, the genitive *-'s* may also be deleted along with *shop* here because it is also extremely light of information content. Note that both these deletions accord with the weight-over-length principle. Incidentally, a similar deletion must be involved in the

derivation of *Parkinson's* and *Alzheimer's* from *Parkinson's disease* and *Alzheimer's disease* respectively.

The following sentence also illustrates a similar deletion of a locative noun triggered by collocational presupposition.

(85) We met *at Bill's (house).*

Note that *at Bill's* collocationally presupposes the locative noun *house* here in such a way that *house* is extremely light of information content. Thus *house* may be deleted here in accordance with the weight-over-length principle. Note incidentally that the locative preposition *at,* along with the genitive *-'s,* plays a key role in collocationally presupposing *house* here. It may be because of this role that *-'s* may not be deleted here while it may be in (84).

The derivation of most nationality/language nouns also involves deletion due to collocational lightening of information. Let us consider the following paraphrase pairs.

(86) a. He speaks *the Korean language.*
b. He speaks ϕ *Korean* ϕ.
(87) a. He's *a Korean citizen.*
b. He's *a Korean* ϕ.

Korean collocationally presupposes *language* in (86a) and *citizen* in (87a). Note that collocationally presupposed nouns here and elsewhere are shared by a number of similar compounds while collocationally presupposing nouns or adjectivals are not thus shared. As a result, presupposed nouns are quasi-suffixal while presupposing nouns or adjectivals are stems of sorts. Thus *language* and *citizen* here are light of information content so that they may be deleted while *Korean* may not, which is in line with the weight-over-length principle. Incidentally, the language noun *Korean* in (86b) does away with *the* because of its status as a firmly established proper noun.

A similar explanation applies to the deletion exemplified by the following sentences.

(88) a. Medicare is for *the aged (people).*
b. We admire *the sublime (thing(s))* in art.

The adjectives here are such that they collocationally presuppose the nouns they premodify. Besides, the nouns are light of inherent semantic content in that they are general-purpose words. Thus they are informationally so light that they may be deleted in accordance with the weight-over-length principle.

The derivation of many de-adjectival nouns can be similarly accounted for. Let us consider the following paraphrase pairs.

(89) a. We crossed *the Rocky Mountains.*
b. We crossed *the Rockies* ϕ.
(90) a. We crossed *the Pacific Ocean.*
b. We crossed *the Pacific* ϕ.
(91) a. He isn't *a homosexual person.*
b. He isn't *a homosexual* ϕ.

Needless to say, the derivation of *the Ohio* from *the Ohio River* goes through the same process although *Ohio* here is not exactly of adjectival origin.[4]

Note at this point that the quasi-suffixal nature of collocationally presupposed nouns throws light on the derivation of the second word in each pair below from the first.

(92) a. smok*ing car* → smok*er*
b. din*ing car* → din*er*
c. wash*ing machine* → wash*er*
d. dry*ing machine* → dry*er*

Note that *-ing* + *car/machine* is quasi-suffixal here and thus light of information content so that it may be shortened to a bona fide suffix, i.e. *-er.* This is all in keeping with the weight-over-length principle.

A current English usage provides us with the following example of deletion due to collocational lightening of information.

(93) *Driving under the influence (of alcohol)* is dangerous.

[4]As we shall see in connection with our discussion of (123) in 2.6., a proper-name-final noun such as *River* here may be deleted only when the proper name is sufficiently familiar to the interlocutors.

Driving under the influence of alcohol is an issue of such universal concern today that it is familiar to most of us. Thus *driving under the influence* has come to collocationally presuppose *of alcohol* for most of us. As a result, *of alcohol* has become informationally so light here that it may be deleted. Once again this is a consequence of the weight-over-length principle.

The notion of collocational presupposition can also help explain the optional deletion in informal English of *is concerned* from *as far as X is concerned,* especially when *X* is long and thus heavy. Note that *as far as* collocationally presupposes *is concerned* here so that *is concerned* is very light of information content. Thus *is concerned* is light enough to get deleted here in accordance with the weight-over-length principle. Note also that *X* is heavier than *is concerned* so that *as far as X is concerned* violates the weight-over-order principle to be discussed in 3.0. Thus the deletion of *is concerned* here has the effect of rendering the expression as a whole more observant of the weight-over-order principle also.

Parts of words may also be deleted under the influence of the weight-over-length principle, as can be seen from the examples below.

(94) a. *be*cause → ϕ 'cause
b. *a*fraid → ϕ 'fraid

Both *be-* and *a-* are light of semantic content in that they are etymologically affixal. Thus they may be deleted in rapid speech, which is in line with the weight-over-length principle.

The derivation of the second word from the first in each pair below involves a similar deletion.

(95) a. *dis*port → ϕ sport
b. *dis*play → ϕ splay
c. *ex*pend → ϕ spend

Note that *dis-* and *ex-* are affixal here so that they are light enough to be deleted. Note also that the last segment of the prefix, i.e. /s/ is not deleted here probably because it serves to distinguish the resulting word from the stem of the source word that often occurs as a free from or from its homonym. That is, /s/ does not get deleted in the derivation of, say, *sport* from *disport* probably because it is needed to set *sport* apart from *port.*

Sometimes a segmental sound may be deleted from a word even though it is not (part of) an affix. Let us consider the two examples below.

(96) a. *A*cadian vs. ϕ Cajun
b. prohibit /prə*h*íbĭt/ vs. prohibition /prə*(h)*ĭbíʃən/

The initial vowel in *Acadian* is atonic, i.e. gets maximally weak stress. Thus it is phonologically light so that it may be deleted and give rise to *Cajun,* a by-form of *Acadian.* Speaking of (96b), (the inherently light) *h* is pre-atonic in *prohibition* and pretonic in *prohibit* so that it is lighter in the former than in the latter. Thus *h* may delete in *prohibition,* but not normally in *prohibit.* Note that both instances of deletion here are in line with the weight-over-length principle.

Note in this connection that we can explain in a principled way why *a* is in free variation with *an* in the following paraphrase pair.

(97) a. *a h*istorical novel
b. *an h*istorical novel

Note that *h* here is pre-atonic so that it is light enough to be optionally deleted. As a result, *h* here may or may not be mute so that we get *a* with the non-mute *h* and *an* with the mute *h.* Thus the weight-over-length principle sheds interesting light on the free variation between the two forms of the indefinite article here.

/h/ is such a light sound that it is often deleted in the *wh*-sequence so that we often end up with free variation between /hw/ and /w/ in words such as the following.

(98) a. when /*hw*en/ or /*w*en/
b. what /*hw*ɑt/ or /*w*ɑt/

The *wh*-sequence here is more likely to retain its /h/ when the *wh*-word in question receives emphasis than when it does not. For example, "What?" is highly likely to retain its /h/ when it expresses a high degree of surprise or dubiety. For the high degree of surprise or dubiety here would render "What?" semantically and thus phonologically heavy.

On the other hand, the *wh*-sequence is highly likely to lose its /h/ when it is pre-atonic, as in the words below.

(99) a. *wh*ĕnévĕr
b. *wh*ătévĕr

Our examples from (96b) on clearly show that a pre-atonic /h/ is more likely to be deleted than a pretonic one. Other consonants also lighten pre-atonically, but they normally stop short of outright deletion on account of their greater inherent weight. For example, /t/ is lighter in *better* than in *deter,* but even the lighter /t/ here does not get deleted. Note that all instances of consonant lightening here are in line with the weight-over-length principle.

It may be in order here to point out that equi-NP deletion and most other types of deletion are also in line with the weight-over-length principle. We have excluded them from consideration here merely on account of space limitations.

2.5. The Derivational Factor

Derivational weight is such that it increases in proportion to the number of derivational stages involved. Derivational weight, thus definable, plays a major role in determining the length of linguistic elements, especially on the lexical level. Let us consider the two word sets below.

(100) a. derive, deriv*ation,* deriv*ational,* deriv*ationally*
b. establish, establish*ment,* establish*mentarian,* establish*mentarianism*

We can see quite clearly here that derivationally heavier words tend to be longer than their derivationally lighter counterparts, which accords with the weight-over-length principle.

Note in this connection that an increment in derivational weight is predicated upon a corresponding increment in semantic weight. Thus semantically heavier words are derivationally heavier also, as can be seen from the following pairs of nouns.

(101) a. command vs. command*ment*
b. laugh vs. laugh*ter*
c. bed vs. bedd*ing*

Note that the second member of each pair here is semantically heavier than the first in that the second member denotes something grander, more comprehensive, or more collective than the first. Thus we can say that deriva-

tional weight is a function of semantic weight.

Assuming that inflection is a kind of derivation, we can explain why the second member of each word pair below is longer than the first.

(102) a. boy vs. boy*s*
b. John vs. John'*s*
c. our vs. our*s*
d. my vs. mi*ne*
e. talk vs. talk*ed*
f. say vs. say*ing*

The second member of each pair here is semantically and thus "derivationally" heavier than the first. Thus the second member here is longer than the first in accordance with the weight-over-length principle.

Note that the following pairs are similarly explainable.

(103) a. th*i*s vs. th*e*se
b. th*a*t vs. th*o*se

The second member of either pair here is longer than the first in that the former has a longer vowel than the latter. This is a consequence of the fact that the second member is semantically and thus "derivationally" heavier than the first.

Inflectional weight sometimes translates into length in such a subtle way that its effect on length is not always readily observable. Let us take the following word pairs for example.

(104) a. s*i*ng vs. s*a*ng
b. s*i*t vs. s*a*t
c. g*e*t vs. g*o*t
d. s*ee* vs. s*aw*

Note that the vowel is lower in the past-tense form than in the present-tense form in each pair here. Since, other things being equal, a lower vowel is longer than a higher vowel, the past-tense form is longer than the present-tense form in each pair above. Note that this is in line with the weight-over-length principle. For the past-tense form is heavier than the present-tense form here in that the former comprises the base form of the verb in question plus the past-tense marker while the latter comprises just the base

form.

Note at this point that an inflectional suffix is normally less conducive to length than is a (bona fide) derivational suffix. As a consequence, the average inflectional suffix is shorter than the average derivational suffix. This is evidently because an inflectional suffix normally carries less semantic weight than a derivational suffix. Note that all this is in line with the weight-over-length principle.

2.6. The Extralinguistic Factor

We have thus far focused on intralinguistic weight as it bears on the length of linguistic elements. As it happens, extralinguistic weight affects linguistic length just as much as does intralinguistic weight. Let us begin our discussion here by considering the following exchange.

(105) A: May I speak to Bob, please?
B: I'm sorry, but he isn't *in* ϕ right now.

The particle *in* here is short for *in the house, in the office,* or the like, as the case may be. Note that *the house, the office,* or the like refers to the immediate context of speech so that it is very light of information content. Thus we may say that *in the house, in the office,* or the like shortens to *in* here as a consequence of the weight-over-length principle.

Expressions of here and now are especially susceptible to deletion due to extralinguistic lightening of information. Let us consider the following paraphrase pairs.

(106) a. It's raining *now.*
b. It's raining ϕ.
(107) a. It's raining *here.*
b. It's raining ϕ.

All the sentences here refer to the rain falling here and now so that *here* and *now* must be present in their deep structures in one form or another. Note here that *here* and *now* refer to the immediate context of speech and are thus light of information content so that they are normally deleted in the process of deep-to-surface mapping. However, they are not deleted when they are weighted with special emphasis such as contrastive emphasis. Thus

normally (106a) contrasts now with some other time while (107a) contrasts here with some other place. On the other hand, (106b) and (107b) imply no such contrast. Thus the use vs. non-use of *here* and *now* in examples such as (106) and (107) above is governed by the weight-over-length principle.

Note in this connecton that the following two sentences constitute a paraphrase pair.

(108) a. The postman hasn't been *here* yet.
b. The postman hasn't been ϕ yet.

The second sentence here arguably derives from the first via *here*-deletion. As already noted, *here* is so closely context-bound that it is light enough to allow deletion, that is, provided that no contrast is intended between here and somewhere else. Thus the derivation of (108b) from (108a) is arguably a consequence of the weight-over-length principle.

It is interesting to note here that time expressions tend to get shorter in proportion to their referential proximity to the present. Let us take the following time expressions for example.

(109) a. today, yesterday, tomorrow
b. the day before yesterday, the day after tomorrow
(110) a. this week, last week, next week
b. the week before last, the week after next

In either (109) or (110), the first set of expressions refers to points in time closer to the present and thus to the immediate context of speech than does the second set. Thus the expressions in the first set are informationally lighter than those in the second set so that the former assume shorter forms than the latter in line with the weight-over-length principle.

Other expressions also tend to get shorter in proportion to their referential proximity to the immediate context of speech. Let us take the following word pairs for example.

(111) a. this vs. that
b. here vs. there

The proximal *this* and *here* are informationally lighter than their respective distal counterparts, i.e. *that* and *there*. Thus the former should be phonologically shorter than the latter if the weight-over-length principle is to be

validated here. This in fact turns out to be the case. For *this* is shorter than *that* in that /is/ is shorter than /æt/ whereas *here* is shorter than *there* in that /hi/ is shorter than /ðɛ/. Note here that /i/, /s/, /h/, and /i/ are shorter than /æ/, /t/, /ð/, and /ɛ/ respectively.

Another interesting example of shortening due to context-bound information lightening has to do with the usages of *B.C.* and *A.D.*, as illustrated by the paraphrase pairs below.

(112) a. She was born in 1960 *B.C.* → *She was born in 1960 ϕ.
b. ?She was born in *A.D.* 1960. → She was born in ϕ 1960.

Note that *B.C.* is obligatory while *A.D.* is normally omitted. In our terms, this is because *B.C.* is referentially much further removed from the present and thus much heavier than is *A.D.* Incidentally, *A.D. 1960* is perfectly acceptable when used, say, in contrast with *1960 B.C.* Needless to say, *A.D.* is justified here because of the additional semantic weight it receives from the contrastive emphasis. *A.D.* is also acceptable when the year in question is very far removed from the present, as in "Mt. Vesuvius erupted in *A.D. 79.*" Note that all these usages are in tune with the weight-over-length principle.

The following paraphrase pairs exemplify still another interesting case of shortening due to context-bound information lightening.

(113) a. I was a student in the *1960*s.
b. I was a student in the *'60*s.
(114) a. I was a student in *1960.*
b. I was a student in *'60.*

The first two digits in *1960* in either pair here refer to the present century and are thus light of information content, which is why they may be deleted. Note that the first two digits may not normally be deleted from the name of any other century than the present one. For they are referentially far removed from the present and thus heavy of information content.

The first two digits may be deleted from the name of a non-present century, however, when the psychological context of speech so dictates. In a discussion of the 16th century, for example, we may refer to *the 1560s* as *the '60s*. To account for cases such as this one, we need a proviso to the effect that the context of speech may also be defined in psychological terms. No matter how the context of speech is defined, the shortening phenomenon

under discussion here is a consequence of the weight-over-length principle.

The apostrophe in *the '60s* and *'60* here is a punctuational trace of sorts for their respective longer forms. Abbreviations often leave similar punctuational traces behind, as in *U.N.* for *United Nations* or in *"The problem is, he is broke"* for *"The problem is that he is broke."* The trace here sometimes takes the form of a function word such as *the,* as in *the Hill* and *the Yard* for *Capitol Hill* and *Scotland Yard* respectively. Phenomena of this sort are dealt with in some detail in Park (in preparation).

Shortening due to context-bound information lightening plays a crucial role in the disambiguation of sentences such as the following.

(115) The ϕ mainland is rich in natural resources.

If a Chinese on Taiwan says this, *the mainland* here is most likely to refer to the Chinese mainland. If a Hawaiian in Honolulu says the same thing, on the other hand, he is most likely to refer to the U.S. mainland. In either case, the gentile adjective suppressed or understood here is the one referring to the mainland maximally close to the immediate context of speech. Thus the adjective suppressed here is light of information content so that the suppression is in line with the weight-over-length principle.

(115) may be differently disambiguated if a psychological context of speech intervenes. In a seminar on Chinese affairs, for example, a Hawaiian in Honolulu may refer to *the Chinese mainland* simply as *the mainland. Chinese* may be deleted here because it refers to the immediate psychological context of speech and is thus light of information content.

Shortening due to context-bound information lightening is also involved in the formation of exophoric pronouns, as can be seen from (116) below.

(116) a. your interlocutor, i.e. the person who is saying this to you here and now → *I*
b. my interlocutor, i.e. the person I am saying this to here and now → *you*
c. the male person that you and I can see here and now → *he*
d. the female person that you and I can see here and now → *she*

The exophoric pronouns on the right-hand side of the arrow are, as it were, short for the longer phrases on the left-hand side of the same. The longer phrases here refer to people in the immediate context of speech so that they are light of information content. Thus they may shorten to their respective

pronominal forms in accordance with the weight-over-length principle.

The so-called ambient *it* is also exophoric and can thus be accounted for along similar lines. Let us take for example the ambient *it,* as used in the sentences below.

(117) a. *It* is too late. (*It*="The time now")
b. *It* is cold today. (*It*="The ambience")

Note that in either sentence here *it* is short for something like the longer phrase given in the parentheses as its gloss. Note further that the longer phrase here refers to one aspect or another of the immediate context of speech. Thus the longer phrase is light of information content so that it may shorten to the ambient *it* in keeping with the weight-over-length principle.

Shortening due to context-bound information lightening also helps explain the gap in the following expression to be found on Coca Cola cans.

(118) Dispose of ϕ properly.

This expression is intended to get Coca Cola consumers to dispose of the Coca Cola can properly after consuming its contents. It is thus short for the longer expression "Dispose of *the can* properly." Note that *the can,* if retained, would be on the can itself so that it would be of minimal weight informationally. It is thus highly redundant and may be deleted in accordance with the weight-over-length principle.

Block language abounds in examples of shortening due to context-bound information lightening. Used mostly in signs and titles, block language comprises phrases such as the following.

(119) a. Exit
b. Keep Out
c. For Sale
d. Das Kapital

All these expressions are short for longer expressions. *Exit,* for one, is short for something like "This is the exit," of which "This is the..." is suppressed because it refers to the immediate context of speech and is thus light of information content. Needless to say, the other phrases in (119) can also be accounted for in a similar way.

Terse commands such as the following also illustrate shortening due to

context-bound information lightening.

(120) a. Out!
b. On your feet!

Note that an underlying verb is understood in each of these commands. This underlying verb is most likely one of locomotion such as *get*, which is light of inherent semantic content. Note that this verb is also closely context-bound and thus gets even lighter of content here so that it may be deleted altogether. We can say exactly the same thing about *"This way!"*, which is arguably short for *"Come this way!"*

Not just verbs of locomotion but other lightweight verbs also are prone to deletion. This is especially the case when the verbs are tightly context-bound, as in the following terse commands.

(121) a. Louder, please.
b. In English, please.
c. One more time, please.

All these commands presuppose an underlying verb of saying or doing such as *say* or *do*, which is light of inherent semantic content. This underlying verb gets its semantic weight further lightened because it is tightly context-bound. Thus the ultimate deletion of this verb here is in line with the weight-over-length principle.

The following sentences illustrate another interesting case of deletion triggered by context-bound information lightening.

(122) a. Please give him my best (*wishes*).
b. What's (*the matter*) with you?

We say "Please give him my best ϕ" only when it is clear from the context that we wish to send the referent of *him* our greetings via the addressee. Thus the expression *wishes*, *regards*, or *greetings* is contextually presupposed so that it is light enough to be deleted. We also use "What's ϕ with you?" only when it is contextually clear that something is wrong or the matter with the referent of *you*. Thus the expression *wrong* or *the matter* is contextually presupposed so that it is light enough to be deleted. Note here that both these instances of deletion are in line with the weight-over-length principle.

General or institutional familiarity also plays a role in shortening due to

context-bound information lightening. Let us take the following paraphrase pair for example.

(123) a. the Philippine *Islands*
b. the Philippines

The Philippine Islands as a geographic/political entity is familiar to most educated people in the world so that it is generally or institutionally familiar to most of us. As a result, *Philippine* collocationally presupposes *Island* so that the latter is very light of information content. Thus *the Philippine Islands* normally shortens to *the Philippines* in accordance with the weight-over-length principle.

Note at this point that a similar shortening is blocked when the referent in question is not yet institutionally familiar enough. For example, *the Taebaek Mountains* may not normally shorten to *the Taebaeks* in an article for a non-Korean readership at first mention at least. For the Korean mountain chain in question is not at all familiar to such a readership so that its name is heavy of information content.

Note in this connection that abbreviations in general presuppose the hearer-reader's institutional familiarity with their referents. Let us take for example the acronym *GU*, as used in the following parpahrase pair.

(124) a. *Georgetown University* overlooks the Potomac river.
b. *GU* overlooks the Potomac River.

Note that *GU* is not normally used unless the hearer-reader is institutionally familiar with its referent. Thus we may say that *Georgetown University* may be abbreviated to GU only when it is quite familiar to the interlocutors so that it is sufficiently light of information content. This is all in keeping with the weight-over-length principle.

It is interesting that overly long names are almost always abbreviated to shorter forms provided that they are reasonably familiar. Let us consider the following paraphrase set.

(125) a. He's from *the Union of Soviet Socialist Republics.*
b. He's from *the Soviet Union.*
c. He's from *the U.S.S.R.*

Although all these three sentences are perfectly gammatical, the first is less natural than the second or the third. Thus *the Union of Soviet Socialist Republics*, which is referentially familiar to us, is almost always abbreviated to either *the Soviet Union* or *the U.S.S.R.* The nearly obligatory abbreviation of *the University of California at Los Angeles* to *UCLA* is amenable to a similar explanation.

3. Weight and Order

3.0. The Weight-Over-Order Principle

Weight plays a pivotal role in determining the order in which linguistic elements occur. The weight-over-order principle given below captures this role of weight as a determinant of order.

> THE HEAVIER A LINGUISTIC ELEMENT IS, THE LATER IT TENDS TO OCCUR; CONVERSELY, THE LIGHTER A LINGUISTIC ELEMENT IS, THE EARLIER IT TENDS TO OCCUR.

Assuming that light elements are generally easier to process than heavy ones, we can say that this principle ensures an easy-to-difficult processing of linguistic signals.[5]

3.1. Adjectivals

Adjectivals are generally heavier of content as postmodifiers than as premodifiers. For postmodifiers tend to carry new information while premodifiers tend to carry given information. It is also generally the case that premodifying adjectivals are lighter of content than the nominal elements they modify while it is the other way around with postmodifying adjectivals. This is all in line with the weight-over-order principle.

Nominal elements do not normally allow premodification when they are (definitely) lighter than the adjectivals that modify them. Let us consider the following paraphrase pairs.

[5]In postpositional languages like Korean, heavy elements often precede light elements so that language processing often proceeds from difficult to easy elements. Which of the two types of processing is more efficient is open to question.

Incidentally, it is in line with the weight-over-order principle that in English governors such as prepositions precede their respective governees.

(126) a. I know *someone nice.*
b. *I know *nice someone.*
(127) a. *Anything new?*
b. **New anything?*

Note that *someone* and *anything* are pronouns while *nice* and *new* are adjectives so that the former are function words while the latter are content words. Thus *someone* and *anything* are (definitely) lighter of content than *nice* and *new* respectively so that the adjectives here must follow the pronouns in accordance with the weight-over-order principle.[6]

Note parenthetically that *something* in the sentence below is only apparently a pronoun.

(128) Here's *a little something* for you.

Something here is a noun in that it has the nominal contentive meaning of "a thing of some value such as a present." Thus it is a content word so that it is fairly heavy of semantic content. Besides, *little* here is light of semantic content so that it comfortably premodifies *something* in accordance with the weight-over-order principle.

The paraphrase pair below illustrates another interesting example of postmodification that can be explained in terms of the weight-over-order principle.

(129) a. This has been the custom here since *time immemorial.*
b. *This has been the custom here since *immemorial time.*

The general-purpose noun *time* is semantically much lighter than *immemorial* so that *time* always precedes *immemorial* in the set phrase *since time immemorial* or *from time immemorial.* Note that this is in line with the weight-over-order principle.

Note in this connection that *immemorial* may not postmodify a noun if the noun is fairly heavy of semantic content. Let us take the following paraphrase pair for example.

[6]Note that *nice* and *new* here modify *-one* and *-thing* respectively, rather than *some-* and *any-*. Thus the postmodification here may serve the additional purpose of minimizing the distance between modifier and modified.

(130) a. *This is *a custom immemorial.*
b. This is *an immemorial custom.*

Custom here is arguably more contentive and thus heavier than *immemorial* so that (130b), but not (130a), is in line with the weight-over-order principle.

When a noun is about the same weight as the adjective that modifies it, it often allows both pre- and postmodification, as in (131) below.

(131) a. In *years past*, they never would have done that.
b. In *past years*, they never would have done that.

Note that neither sentence here actively violates the weight-over-order principle so that both are in line therewith.

Note here that *past* may not postmodify a noun if it is definitely lighter than the noun. Thus (132b) below, but not (132a), is grammatical.

(132) a. *His *successes past* are undeniable.
b. His *past successes* are undeniable.

Successes here is definitely heavier than *past* so that (132b), but not (132a), is in line with the weight-over-order principle.

There is often a semantic difference between pre- and postmodification when both are allowed, as in (133) below.

(133) a. *Newburg* lobster (=lobster in/from Newburg)
b. lobster *Newburg* (=lobster à la Newburg)

Note that *Newburg* is semantically heavier in (133b) than in (133a). Note also that *Newburg* is heavier of content than *lobster* in (133b) while it is the other way around in (133a). Thus both expressions here follow the weight-over-order principle.

The difference in word order between the two expressions below can be explained along similar lines.

(134) a. *proper* ethics (=correct ethics)
b. ethics *proper* (=ethics in the strict sense of the word)

The following pair of noun phrases may also be similarly accounted for.

(135) a. *Korean* things (=products of Korea)
b. things *Korean* (=things characteristic of Korea)

The general-purpose noun *things* is arguably heavier in *Korean things* than in *things Korean* in that it is referentially more concrete in the former than in the latter. Besides, *Korean* here is semantically heavier in *things Korean* than in *Korean things*. Furthermore, *things* is arguably heavier than *Korean* in *Korean things* while it is the other way around in *things Korean*. Thus both noun phrases here are in line with the weight-over-order principle.

The weight-over-order principle is also apparently instrumental in explaining the historical development of (136a) below into (136b).

(136) a. *martial* court
b. court *martial*

Martial is fairly heavy of inherent semantic content so that it is at least as heavy as, if not heavier than, *court*. Besides, the original term *martial court* was used in contradistinction to (*civilian*) *court* so that *martial* received considerable contrastive emphasis. As a result, *martial* became much heavier than *court* quite early on so that they ended up trading places, giving rise to the contemporary term *court martial*. Note that this change took place under the pressure of the weight-over-order principle.

The weight-over-order principle can also help explain why *apparent* is postpositive in the following expression.

(137) heir *apparent*

Note that *apparent* is far heavier of semantic content in *heir apparent* than in *apparent heir*. Besides, *heir apparent* contrasts with *heir presumptive* so that in (137) *apparent* gets contrastive stress and is thus much heavier than *heir*. Thus *apparent* postmodifies *heir* in (137) in accordance with the weight-over-order principle.

The weight-over-order principle also throws light on the postpositive adjectivals in expressions such as the following.

(138) a. Ivan *the Terrible*
b. John F. Kennedy, *Jr.*
c. John *the Baptist*

The adjectivals here serve to set apart the bearers of the names in question from all other persons of the same names. Thus the adjectivals receive contrastive emphasis so that they are semantically heavier than the names they modify. As a result, the adjectivals here are postpositive in accordance with the weight-over-order principle.

The following paraphrase pairs also show that adjectivals pre- or postmodify nouns, depending on their weight relative to the nouns they modify.

(139) a. *Rockefeller* University
b. *the University *of Rockefeller*
(140) a. ?*William and Mary* College
b. the College *of William and Mary*

Note that human nouns are generally lighter of information content than non-human nouns in that the former are referentially closer to us humans than the latter. Thus *Rockefeller University* follows the weight-over-order principle while *the University of Rockefeller* does not. Note that even the addition of *of* to *Rockefeller* does not render the adjectival here heavier than the noun it modifies, which apparently indicates that human nouns are very light of content.

Note that *William and Mary*, although human in reference, is rather heavy of content in that it is a coordinated noun phrase. In fact, it is slightly heavier than *college* here so that (140a) is not quite in line with the weight-over-order principle while (140b) is perfectly in line therewith.

Note in this connection that (141) below is similar in nature to (139) above.

(141) a. *John's* head
b. *the head *of John*

The human noun *John* is informationally lighter than the non-human noun *head* so that *John's head* follows the weight-over-order principle while *the head of John* does not.

Suppose that we replaced *John* with *John the Baptist* in (141b). Then we would get the following phrase, which is perfectly grammatical.

(142) the head *of John the Baptist*

Note that *the Baptist* adds considerably to the weight of *John* so that (142), unlike (141b), complies with the weight-over-order principle.

Note at this point that clausal modifiers normally postmodify nouns because clauses are structurally heavier than nouns. Incidentally, we count prepositional phrases as clauses for our purposes here because prepositions and their complements are arguably verbs and their objects respectively. Thus we may say that *asleep*, *galore*, and other adjectives that originate in prepositional phrases normally postmodify nouns because they are structurally heavier than the nouns.

3.2. Adverbials

When two or more adverbials co-occur, they tend to occur in ascending order of weight, which is exactly the opposite of what happens in postpositional languages like Korean. Let us consider the following paraphrase pairs.

(143) a. I was born in *Seoul, Korea.*
b. *I was born in *Korea, Seoul.*
(144) a. I was born in *January, 1960.*
b. *I was born in *1960, January.*

Note that *Korea* and *1960* here are referentially larger and thus semantically heavier than *Seoul* and *January* respectively. Thus the grammatical sentences here are those that follow the weight-over-order principle while the ungrammatical ones are those that do not.

The weight-over-order principle can also help explain why (145a) below, but not (145b), is grammatical.

(145) a. I did that *every day last week.*
b. *I did that *last week every day.*

A week is a larger unit of time than a day so that (145a), but not (145b), is in line with the weight-over-order principle, which is why the former is grammatical while the latter is not.

It is interesting that the weight-over-order principle also throws light on why (146a) below, but not (146b), is grammatical.

(146) a. Today is *Tuesday, May 2.*
b. *Today is *May 2, Tuesday.*

A month is superordinate to a week so that a day of the month is referentially larger and thus semantically heavier than a day of the week. Thus *May 2* here is semantically (as well as structurally) heavier than *Tuesday* so that (146a), but not (146b), complies with the weight-over-order principle, which is why the former, but not the latter, is grammatical. Incidentally, *May 2* puts more weight on *2* than on *May* while *2 May* puts more weight on *May* than on *2* so that both expressions comply with the weight-over-order principle.

Note at this point that, other things being equal, a time adverbial is referentially superordinate to, and thus semantically heavier than, a place adverbial. This is why the first sentence below is grammatical while the second is not.

(147) a. I was born *in Korea in 1960.*
b. *I was born *in 1960 in Korea.*

3.3. Subjects vs. Predicates

The average English sentence begins with a subject that is light of content. It then proceeds to a predicate that is heavy of content. Thus it proceeds from a typically light subject to a typically heavy predicate, which is in line with the weight-over-order principle.

Reflective of this general preference for light subjects is the frequent occurrence in the subject slot of human noun phrases, including pronouns and proper nouns. Since human noun phrases are referentially more familiar to us and thus informationally lighter than non-human noun phrases, this is apparently in line with the weight-over-order principle. Note in this connection that the subject choice hierarchy of Fillmore (1968) also arguably follows from the weight-over-order principle. Rephrased in our terms, the hierarchy says roughly: "The more referentially human a noun phrase is, the more likely it is to fill the subject slot in an active sentence." We assume here that the agentive case role is referentially more human than the instrumental case role, which in turn is more so than the objective case role.

Heavy subjects are maximally avoided because they pose something of a problem for the weight-over-order principle. Let us consider the following

paraphrase pair.

(148) a. (?) *A book* is on the table.
b. *There's a book* on the table.

The indefinite noun phrase *a book* is informationally heavy in that it carries new information. The verb phrase *is on the table* is a light predicate in that *is* is a lightweight verb and that the definite noun phrase *the table* carries old information. Thus the subject is slightly heavier of information content than the predicate in (148a), which is out of line with the weight-over-order principle.

Note that (148a) is often repaired to (148b) by shunting *a book* in back of *is* and filling the subject slot with the existential *there,* which is virtually empty of meaning. Thus in (148b) the subject is much lighter than the predicate, which is in line with the weight-over-order principle.

Informationally heavy indefinite subject noun phrases may also be avoided in other ways, one of which is illustrated by the following paraphrase pair.

(149) a. (?)*A lake* is near *my cottage.*
b. *My cottage* is near *a lake.*

Note that the indefinite *a lake* is informationally heavier than the definite *my cottage* so that (149b) is more in line with the weight-over-order principle than (149a). Thus (149b) is stylistically superior to (149a) so that we normally repair (149a) to (149b), thus avoiding the heavy indefinite noun phrase in the subject slot.

Note that clausal subjects are normally avoided because they are too heavy to occupy the subject slot. Let us consider the following paraphrase set.

(150) a. *They* said that he was a genius.
b. **That he was a genius* was said.
c. *It* was said that he was a genius.

Ordinarily, (150b) should be the passive version of (150a) except that it is ungrammatical and thus avoided. Note that (150b) violates the weight-over-order principle in that its clausal subject *That he was a genius* is much too heavy, in fact, far heavier than its predicate *was said.* Thus (150b) is often

repaired to (150c) by means of *it*-extraposition. Note that (150c) is in line with the weight-over-order principle in that its subject *it* is far lighter than its predicate *was said that he was a genius*. Incidentally, (150b) may also be repaired to "He was said to be a genius," which also complies with the weight-over-order principle.

The predicate in (150b) is light because the general-purpose verb *say* is light of inherent semantic content. By replacing *said* with *reported* and *confirmed* in (150b), we get the following sentences.

(151) a. ?That he was a genius *was reported*.
b. (?)That he was a genius *was confirmed*.

Note here that *confirm* is slightly heavier than *report*, which in turn is slightly heavier than *say*. Thus (151b) is more in line with the weight-over-order principle than (151a), which in turn is more so than (150b). This is why (151b) is more natural than (151a), which in turn is more natural than (150b). Note here that (151b), though the most natural of the three sentences in question, is still slightly odd. This is because the predicate *was confirmed* is still slightly lighter than the subject *That he was a genius*, which is out of line with the weight-over-order principle. Suppose here that *not* were used to negate the predicate in either (151a) or (151b). Then the predicate would gain considerable weight from this *not* so that neither sentence here would violate the weight-over-order principle any more. Thus both "That he was a genius was *not* reported" and "That he was a genius was *not* confirmed" would be perfectly natural.

Note that *it*-extraposition may also be used to repair the first sentence in each pair below to the second.

(152) a. **That everything is fine* seems.
b. *It* seems that everything is fine.
(153) a. **That everything is fine* appears.
b. *It* appears that everything is fine.
(154) a. **That everything is fine* turns out.
b. *It* turns out that everything is fine.

The ungrammatical first sentence in each pair here violates the weight-over-order principle in that the subject is much heavier than the predicate. We may repair this first sentence to the second by shunting the problematic clausal subject back and putting in its place the expletive *it*, which is virtual-

ly empty of meaning. Note that the sentence thus repaired is fully in line with the weight-over-order principle. For the subject is now far lighter than the predicate.

The starred sentences in (152), (153), and (154) may be repaired in a different way. We may weight the light predicate here with (*to be*) *true* or (*to be*) *the case,* as in the following repaired versions of (152a).

(155) a. That everything is fine *seems (to be) true.*
b. That everything is fine *seems (to be) the case.*

Note that (*to be*) *true* or (*to be*) *the case* adds substantial structural weight to the predicate here. On account of this added weight, either reparied version here is in line with the weight-over-order principle, which is why it is perfectly natural.

It may very well be the case that either (155a) or (155b) is the ultimate source for (152b). Assuming that this is indeed the case, we may say that (*to be*) *true* or (*to be*) *the case* may be deleted only after *it*-extraposition has applied to (155). If we should follow the derivational route suggested here, then the starred (152a), (153a), and (154a) do not have to be generated in the first place. Note that the derivational route suggested here allows the deletion of (*to be*) *true* or (*to be*) *the case* only when it is embedded in an unusually long predicate. Thus the deletion here arguably serves the purpose of preventing the predicate in question from becoming overly heavy.

The kind of derivational route suggested in the preceding paragraph makes a great deal of sense in that it also throws light on data such as the following.

(156) a. That he is to blame is not *the case.*
b. *That he is to blame is not ϕ.
c. It is not (*the case*) that he is to blame.

We may replace *is not* here with *may be* or *could be* without any change in grammaticality judgment.

Note here that the derivation of the adverbial *maybe* or *could be,* as exemplified below, also apparently involves a similar deletion of *the case* or *true.*

(157) A: Do you think he's to blame?
B: *Maybe./Could be.*

Note that *Maybe* and *Could be* here may ultimately derive from *That may be the case* and *That could be the case* respectively. Note further that *the case* here is sentence-final and may still be deleted. This is apparently because the subject itself is light enough to be deleted, in fact, lighter than what is left of the predicate after the deletion of *the case*. In other words, the sentence-final *the case* may be deleted here apparently because it does not lead to a violation of the weight-over-order principle.

Note at this point that the following sentence is a perfectly grammatical paraphrase of the ungrammatical (153a).

(158) That everything is fine *is apparent.*

Note that *is apparent* is structurally much heavier than *appears* although both are about the same semantic weight. Thus (158) is in line with the weight-over-order principle while (153a) is not.

Our discussion here clearly demonstrates that English is intolerant of overly light predicates, especially when they are lighter than the subjects. English is, in fact, intolerant of overly light predicates even when they are not necessarily lighter than subjects. Let us take the following paraphrase pairs for example.

(159) a. *He *suggested.*
b. He *made a suggestion.*
(160) a. ?He *sat.*
b. He *sat down.*
(161) a. ?He *stood.*
b. He *stood up.*
(162) a. (?)*Bathe.*
b. *Take a bath.*
(163) a. (?)*Sip.*
b. *Have a sip.*
(164) a. (?)Let's *shop.*
b. Let's *go shopping.*
(165) a. ?Butter *cuts.*
b. Butter *cuts easily.*

Note that the predicate in the first sentence of each pair here is structurally very light in that it comprises just one relatively lightweight verb. On the other hand, the predicate in the second sentence of the same pair is heavier

by one or two words. Thus the first sentence is, as it were, less in line with the weight-over-order principle than the second, which is why the former is less natural than the latter.

The desire to avoid an excessively light prediate may underlie the use of the expletive *it* with many zero-derived denominal intransitive verbs such as those used in the following sentences.

(166) a. He'd probably take fright and *leg it.*
b. He decided to *foot it,* rather than wait for the next bus.

3.4. Information Focus

The word that receives the information focus of a tone unit is informationally heavier than any other word in the same tone unit. This focused word normally comes at the end of the tone unit in question, which is in line with the weight-over-order principle. Let us consider the following paraphrase pair.

(167) a. John blamed *Mary* for *the accident.*
b. John blamed *the accident* on *Mary.*

(167a) and (167b) would be appropriate as responses to "What did John blame Mary for?" and "Who did John blame the accident on?" respectively. Thus the information focus is on *the accident* in (167a) and on *Mary* in (167b) so that the two noun phrases are heavier than any other phrase in their respective sentences. Note here that *Mary* and *the accident* occur in ascending order of weight in either sentence here. Note also that the focused noun phrase is sentence-final in either sentence. All this accords with the weight-over-order principle.

It is interesting to note here that the focused noun phrase in an active sentence may not become the passive subject. Thus (167) may passivize into (168) below, but not into (169).

(168) a. *Mary* was blamed for *the accident.*
b. *The accident* was blamed on *Mary.*
(169) a. **The accident* was blamed *Mary* for.
b. **Mary* was blamed *the accident* on.

In (168), the focused noun phrase in the active (source) sentence is sentence-final and comes later than the unfocused noun phrase in the active (source) sentence. In (169), on the other hand, the focused noun phrase not just is not sentence-final but also comes earlier than the unfocused noun phrase. Thus the passives in (168) follow the weight-over-order principle while those in (169) do not, which is why the former only are grammatical.

Information focus is also relevant to a discussion of the difference in word order between the two sentences below.

(170) a. She gave *John a book.*
b. She gave *a book* to *John.*

The information focus is on *a book* in (170a) and on *John* in (170b) so that the two noun phrases are heavier than any other phrase in their respective sentences. Note here that *John* and *a book* occur in ascending order of weight in either sentence and also that the focused noun phrase is sentence-final in either sentence. This is, of course, in compliance with the weight-over-order principle.

Note in this connection that, prior to dative movement, the indirect object slot is unfocused and thus light so that it favors (inherently light) human noun phrases over (inherently heavy) non-human noun phrases. Let us consider the following paraphrase pairs.

(171) a. Save *Mrs. Smith* a place.
b. Save a place for *Mrs. Smith.*
(172) a. ?Save *the tree* a place.
b. Save a place for *the tree.*

The indirect-object slot is filled by a human noun phrase in (171a) and by a non-human noun phrase in (172a) so that (171a) is more in line with the weight-over-order principle than (172a). We may say here that dative movement applies optionally to (171a) and obligatorily to (172a), that is, on the assumption that the second sentence derives from the first in either pair above. This is also in line with the weight-over-order principle.

Note here that (170) may passivize into (173) below, but not into (174), for exactly the same reason that (167) may passivize into (168), but not into (169).

(173) a. *John* was given *a book*.
b. *A book* was given to *John*.
(174) a. ?*A book* was given *John*.
b. *?*John* was given *a book* to.

On the basis of our discussion in the preceding parapgraphs, we may propose a constraint on passivization to the effect that the focused noun phrase in an active sentence may not become subject in its passive version. It is significant that this focused noun-phrase constraint follows from the weight-over-order principle.

The weight-over-order principle can also help explain why the agent *by*-phrase is absent from most passive sentences. Note that the agent *by*-phrase originates in the active subject, which is normally light of information content. Suppose that this light active subject were retained in a passive sentence in the form of an agent *by*-phrase. Then it would have to come at the very end of the passive sentence, a position typically reserved for a heavy element such as the focused noun phrase of the sentence. Thus an agent *by*-phrase would normally violate the weight-over-order principle, which is apparently the main reason why most passives do away with it.

Note in this connection that the agent *by*-phrase is perfectly natural in the following exchange.

(175) A: *Who* wrote this essay?
B: Oh, it was written *by John*.

The question in this exchange is such that the information focus in the answer is on the writer of the essay, i.e. on the active subject of *write*. Being the informationally heaviest element in the answer, this active subject should ideally take sentence-final position in one way or another. Here this sentence-final position is achieved by means of passivization, which puts the focused active subject at the tail end of the sentence in the form of an agent *by*-phrase. Note that the use of the agent *by*-phrase here is in line with the weight-over-order principle, not in violation thereof.

It is interesting that an active subject is sometimes sentence-final, as in the following exchange.

(176) A: Says *who?*
B: Says *me.*

Note that both *who* and *me* are subjects here so that they should normally be sentence-initial. However, they are focused upon and thus heavier than (the inherently light) *says* in both sentences so that they are both sentence-final in accordance with the weight-over-order principle.

Note here that the object of *says* is deleted in (176). We may speculate here that the subject is moved back of the verb in both sentences of (176) to help fill the vacuum left by the deleted object. Thus we may say that the movement of the subject to post-verbal position here is probably triggered at least in part by the deletion of the object. As it turns out, deletion often appears to trigger inversion. We will come back to deletion as a trigger for inversion in connection with our discussion of (191) through (196) as well as (218) and (225).

Information focus also plays a crucial role in fixing the apparently abnormal word order in the following proverbs.

(177) a. He that mischief *hatches,* mischief *catches.*
b. Early *sow,* early *mow.*
c. First *come,* first *served.*
d. Nothing *seek,* nothing *find.*
e. Nothing *crave,* nothing *have.*

Note that each proverb here comprises two clauses and that the two clauses terminate with words that contrast with each other. Thus the clause-final words here receive information focus through contrastive emphasis with the result that they are the heaviest words in their respective clauses. Seen in this light, the clause-finality of the italicized words in (177) is a consequence of the weight-over-order principle. We may also note in passing here that rhyme may be relevant to a discussion of the clause-finality of the focused words in the first two proverbs of (177). Considering that *have* used to rhyme with *crave,* we may also argue that rhyme is relevant to the clause-finality of the italicized words in the last proverb above.

A similar explanation is applicable to the clause-finality of *I* in the following sentences.

(178) a. *John* is hungry, and so'm *I.*
b. *John* isn't hungry, and neither'm *I.*

I here contrasts with *John* so that it receives information focus. Thus *I* is the heaviest word in the clause in which it occurs so that it is clause-final in accordance with the weight-over-order principle. *John*, on the other hand, is clause-initial because it bears old information and is thus light.

3.5. Flipflop Words

A flipflop word is a lightweight word that varies its position according to the weight of a word in its immediate vicinity. *All*, as used in the paraphrase pairs below, is a case in point.

(179) a. They *can all* come and see us.
b. ?They *all can* come and see us.
(180) a. *They *hope all* to come and see us.
b. They *all hope* to come and see us.

Note that *all* is slightly heavier than *can* while it is lighter than *hope*. Thus the natural sentences here abide by the weight-over-order principle with respect to the two words in italics while the odd sentences do not.

The variable position of *all*, as illustrated by the paraphrase pairs below, also accords with the weight-over-order principle.

(181) a. They told *us all* to wait.
b. *They told *all us* to wait.
(182) a. *They told *the men all* to wait.
b. They told *all the men* to wait.

All is heavier than *us* while it is lighter than *the men* so that the grammatical sentences here follow the weight-over-order principle while the ungrammatical ones do not.

Note at this point that we can repair (181b) to (183) below.

(183) They told *all of us* to wait.

Note that *all of us* comprises two immediate constituents, i.e. *all* and *of us*, of which *all* is structurally lighter than *of us*. Thus *all of us*, unlike *all us*, is perfectly in line with the weight-over-order principle, which is why (183) is grammatical.

Each and *both* are flipflop words of exactly the same kind as *all* so that they all behave identically. Thus we may replace *all* with either *each* or *both* in (179) through (183) and still get the same result.

Frequency adverbs and the negative adverb *not* are also flipflop words. They vary their position according to the weight of the first word in the verb phrase, as can be seen from the two sets of sentences below.

(184) a. He *is always* late.
b. He *can always* walk.
c. He *always walks.*
(185) a. He *is not* late.
b. He *cannot* walk.
c. He *does not walk.*

Note that *always* and *not* are semantically heavier than *is* and *can* (as well as *does*) while they are semantically lighter than *walk(s)*. Thus *always* and *not* follow *is* and *can* (as well as *does*) while they precede *walk(s)*, which is in line with the weight-over-order principle.

Note that *is* and *can* may gain weight through contrastive or affirmative emphasis and thus be heavier than *always*, in which case *is* and *can* must follow *always*. This is also in line with the weight-over-order principle.

Note in this connection that *always* precedes *is* and *can* in B's responses below in apparent contradiction of the weight-over-order principle.

(186) A: Is he always late?
B: Yes, he *always is*. (=*?Yes, he *is always.*)
(187) A: Can he always walk?
B: Yes, he *always can*. (=*?Yes, he *can always.*)

Note that the clause-final *is* and *can* here are short for, and thus carry the weight of, *is late* and *can walk* respectively. Thus they are heavier than *always* so that they follow *always* in both responses above, not in violation of the weight-over-order principle but in compliance therewith.

The flipflop word *not* is about the same weight as a number of lightweight verbs so that it may either precede or follow them. The following paraphrase pairs illustrate the positional variability of *not* when it is in construction with three such verbs.

(188) a. He *has not* a farm.
b. He *does not have* a farm.
(189) a. He *need not* go.
b. He *does not need* to go.
(190) a. He *used not* to visit us.
b. He *did not use* to visit us.

Note that two elements of similar weight may apparently be ordered relative to each other in either of two possible ways. For either order does not actively violate the weight-over-order principle. This is apparently why both members of each pair above are about equally in line with the weight-over-order principle and thus about equally natural.

Note that (189a), if put in the past tense, would sound odd. For the past tense would add its weight to *need*, which is itself fairly heavy, so that *needed* would be far heavier than *not*. Incidentally, (190a) may not be as natural as (190b) in that *used* may be heavier than *not* in the same way that *needed* is far heavier than *not*. Note also that the dynamic *have*, as in *have fun*, is heavy enough to normally require *do*-support in negation (and in question-formation as well). This is all in line with the weight-over-order principle.

3.6. Inversion

The weight-over-order principle sheds plenty of light on the phenomenon of inversion. Let us begin our discussion here with the following paraphrase pair.

(191) a. *When summer comes*, I'll go to Chejudo.
b. *Come summer*, I'll go to Chejudo.

Details aside, (191b) arguably derives from (191a) via conjunction deletion and subject-verb inversion. Note that the lightweight verb *come* is lighter than *summer* here so that the inversion involved here is in line with the weight-over-order principle.

Perhaps similarly explainable is the derivation of the subjunctive *come* in the idiomatic expression *come what may*. Let us consider the paraphrase pair below.

(192) a. *No matter what may come*, I will stay with you.
b. *Come what may*, I will stay with you.

Here again the second sentence arguably derives from the first via conjunction deletion and "subject-verb" inversion. Note that *come* is lighter than *what may* so that they change places in the course of the derivation here in accordance with the weight-over-order principle.

Note here that in both (191) and (192) subject-verb inversion occurs in tandem with conjunction deletion. The verb may move forward in the inversion here to fill the vacuum left by the conjunction deleted so that the deletion arguably triggers the inversion. These remarks may also apply to subject-verb inversion exemplified by (193) and (195) below.

(193) a. *If he were* not rich, he wouldn't buy a BMW.
b. *Were he* not rich, he wouldn't buy a BMW.
(194) a. *If he weren't* rich, he wouldn't buy a BMW.
b. **Weren't he* rich, he wouldn't buy a BMW.
(195) a. *If you should* change your mind, please let me know.
b. *Should you change* your mind, please let me know.
(196) a. *If I knew*, I'd tell you.
b. **Knew I*, I'd tell you.

The second sentence in each pair here arguably derives from the first via conjunction deletion and subject-verb inversion. Note that the subject-verb inversion here, along with the rest of the derivational process, is allowed only when the first word in the verb phrase in question is about as light as, or lighter than, the subject. Thus the inversion phenomenon illustrated here is in line with the weight-over-order principle.

The weight-over-order principle can also help explain "the subject-verb" inversion exemplified by the following sentence.

(197) He said *he'd win,* and *win he did.*

The second token of *win* here is a mere repetition of the first so that it is virtually empty of information content. On the other hand, *he did* is more than a mere repetition of *he'd* so that it is heavier than the second token of *win.* Thus *win he did,* which arguably derives from *he did win* via subject-verb inversion, is in line with the weight-over-order principle.

Subject-verb inversion involved in question formation is also predicated on the levity of the first word in the verb phrase. Let us consider the following statement-question pairs.

(198) a. *He is* late.
b. *Is he* late?
(199) a. *He can* walk.
b. *Can he* walk?
(200) a. *He* walk*s*.
b. *Does he* walk?

The second sentence in each pair here is supposed to derive from the first via subject-verb inversion. In (198) and (199), the first word in the verb phrase is light enough to trade places with the subject so that the inversion involved is quite straightforward. In (200), on the other hand, the first word in the verb phrase, i.e. *walks,* is too heavy to change places with the subject so that the inversion here is not quite as straightforward. Note here that the inflectional suffix *-s* on *walks* is on a par with an auxiliary verb. Thus we may argue that this *-s* is lighter than *he walk-* here so that they may trade places, with *-s* materializing as *does* when fronted (to sentence-initial position). Seen in this light, the subject-verb inversion illustrated by (198) through (200) is all in line with the weight-over-order principle.

With a few verbs of ambivalent weight, the inversion involved in question formation may or may not be straightforward. Let us consider the following paraphrase pairs.

(201) a. *Has he* a farm?
b. *Does he have* a farm?
(202) a. *Need he* go?
b. *Does he need* to go?

The verbs in question here are of light, ambivalent weight in that they can be on either side of the borderline between auxiliary and lexical verbs. As a result, they may or may not trade places with the subject in the process of question formation. Either way, the weight-over-order principle is not violated so that both members of either pair here are grammatical.

It is interesting that we get the following pairs by putting the verbs in (201) and (202) in the past tense.

(203) a. (?)*Had he* a farm?
b. *Did he have* a farm?
(204) a. **Needed he* go?
b. *Did he need* to go?

The past tense evidently adds to the weight of the verb in either pair here so that a straightforward subject-verb inversion here runs afoul of the weight-over-order principle. *Need* is an inherently heavier verb than *have,* which may explain why (204a) is far less natural than (203a).

Note at this point that *Need* in (202a) is bare of the usual third-person-present-singular marker *-s*. This may be because *Need* is already fairly heavy so that *-s* is jettisoned, so to speak, so as to make the sentence better comply with the weight-over-order principle. Note in this connection that the third-person-present-singular *-s* is obligatorily jettisoned from operators such as *can, may must,* and *will*. This is probably because an operator is the first element in a verb phrase and is thus often involved in subject-verb inversion so that it must watch its weight. It may be for the same reason that an operator normally lacks a true past-tense marker. If correct, this is all in line with the weight-over-order principle.

Our discussion here also throws light on why (205b) below is more natural than (205a).

(205) a. (?)*Used he* to live in Seoul?
b. *Did he use* to live in Seoul?

Used here originates in a lexical verb and is in the past tense so that it is apparently a bit too heavy to trade places with the subject *he*. This is evidently why (205a) is slightly less natural than (205b).

The general-purpose verb *say* apparently used to be of ambivalent weight until not too long ago. Let us consider the following paraphrase pair.

(206) a. What *say you* to that?
b. What *do you say* to that?

Sentences like (206a) used to be, and may in some dialects still be, interchangeable with sentences like (206b). This evidently means that *say* was until recently slightly lighter than it is today.

Something similar may be said about the general-purpose verbs *come* and *go* in light of the following expressions, which may be slightly archaic today.

(207) a. How *goes the enemy?* (=What time is it?)
b. How *goes it* with you?
(208) a. How *comes it* that you know everything so well?
b. How *came you* to be up so early?

Note that *come* and *go* precede the subject in each question here, which is an apparent result of subject-verb inversion. Although this usage has lost much of its currency, it apparently indicates that *come* and *go* must have until recently been almost as light as auxiliary verbs.

From our discussion up to this point, we can see that subject-verb inversion can normally take place only when the first word in the verb phrase is lighter than, or about as light as, the subject. The following paraphrase pairs show that subject-verb inversion is disallowed when the subject is distinctly lighter than the first word in the verb phrase.

(209) a. I'm as hungry as *John is.*
b. I'm as hungry as *is John.*
(210) a. I'm as hungry as *he is.*
b. *I'm as hungry as *is he.*

Note that *is* here is short for *is hungry* so that it is fairly heavy of semantic content. It is apparently about as heavy as *John* and distinctly heavier than *he* so that (210b) violates the weight-over-order principle while (209b) does not. Needless to say, this is why subject-verb inversion may apply to (209a), but not to (210a).

The following paraphrase pairs show that subject-verb inversion is better avoided even when the subject is only slightly lighter than the first word in the verb phrase.

(211) a. "Good!" *Bill said.*
b. "Good!" *said Bill.*
(212) a. "Good!" *he said.*
b. (?)"Good!" *said he.*

Note that *said* is about as heavy as *Bill* here while it is slightly heavier than *he.* Thus (212b) is not quite in line with the weight-over-order principle while (211b) is, which is why (212b) is stylistically suspect and is thus better avoided.

Note that subject-verb inversion is obligatory in the derivation of (213a) below while it is disallowed in the derivation of (213b).

(213) a. Here *comes John.*
b. Here *he comes.*

Note that the subject is heavier than the first word in the verb phrase in (213a) while it is the other way around in (213b). Thus both sentences here are in line with the weight-over-order principle. Note that the weight-over-order principle determines whether or not subject-verb inversion should apply here.

The following paraphrase pairs illustrate what may be called "concessive" subject-complement inversion, which is also explainable in terms of the weight-over-order principle.

(214) a. *Though he was clever,* he couldn't outwit her.
b. *Clever though he was,* he couldn't outwit her.
(215) a. *Though he was a genius,* he couldn't outwit her.
b. *Genius though he was,* he couldn't outwit her.

Clever and *genius* are fronted here only when the cleverness or the geniusness of the person in question is highly given and thus taken for granted. In fact, they are fronted only when they are so light as to be lighter than what precedes them, which accords with the weight-over-order principle. Note incidentally that the indefinite article *a* is deleted when *genius* gets fronted. This apparently follows from the fact that the geniusness of the person in question is so given here that *a*, a bearer of new information, is out of place here.

Pro-forms often figure in inversion. They are typically so light that they often get fronted as part of an inversion process, as is the case with *so* and *neither* in the sentences below.

(216) a. Bill is here, and *so* is John.
b. Bill isn't here, and *neither* is John.

The pro-forms *so* and *neither* are so light that they get fronted to clause-initial position. Incidentally, *John* is clause-final because it is contrastively focused upon and thus heavier than any other word in the clause. This is, of course, all in line with the weight-over-order principle.

The pro-form *so* is clause-initial in the following exchange for the same reason that it is clause-initial in (216a).

(217) A: This is good.
B: *So* it is.

Note that *it* here, unlike *John* in (216a), is not focused upon and that it is a mere pro-form. Thus it is extremely light of information content so that it is not clause-final here.

The pro-form *so* may also figure in an optional inversion, as is illustrated by the following paraphrase pair.

(218) a. In *doing so*, you hurt her feelings.
b. In *so doing*, you hurt her feelings.

The pro-form *so* is apparently about as light as the general-purpose verb *do* so that the weight-over-order principle may apply in either direction here. Thus both *doing so* and *so doing* are allowed here so that the inversion here is optional. Note that the logical subject of *doing* in (218), which is *you,* is missing from the surface and that the inversion under discussion here may be designed to fill the vacuum left by this missing subject of *doing*. The same is evidently true of the inversion involved in the derivation of "*So* to *do,* he needs your help" from "To *do so,* he needs your help."

The following paraphrase pair exemplifies the fronting of the pro-form *such* in a similar inversion triggered by the weight-over-order principle.

(219) a. ?Life is *such*.
b. *Such* is life.

The pro-form *such* is much lighter than *life* here so that (219b) complies with the weight-over-order principle while (219a) does not, which is why (219b) is natural while (219a) is odd. Note that replacing *such* with *this* or *that* here yields about the same result.

Let us now turn our attention to particle-object inversion, as exemplified by the following paraphrase pair.

(220) a. He mixed *up the twins*.
b. He mixed *the twins up*.

Note that the information focus is on *the twins* in (220a) and on *up* in (220b). Assuming that (220b) derives from something like (220a), we may say that the information focus on the particle *up* triggers the particle-object inversion involved in the derivation here. Note that this is all in line with the weight-over-order principle.

It is interesting that particle-object inversion is obligatory when the object

is inherently lighter than the particle. Let us consider the two paraphrase pairs below.

(221) a. *He mixed *up them.*
b. He mixed *them up.*
(222) a. ??He mixed *up things.*
b. He mixed *things up.*

Note that the pronoun *them* and the general-purpose word *things* are inherently much lighter than the particle *up* here. As a result, the first sentence in either pair violates the weight-over-order principle while the second does not. Assuming, as we did in connection with (220), that the second sentence derives from something like the first in either pair here, we can see that particle-object inversion is obligatory here.

The following paraphrase pair illustrates another instance of particle-object inversion.

(223) a. ??They turned *on and off the heat.*
b. They turned *the heat on and off.*

The coordinate particle phrase here is considerably heavier than the object noun phrase so that (223a) violates the weight-over-order principle while (223b) does not. Thus (223a), which is definitely odd, is better repaired to (223b) by means of particle-object inversion.

There are cases that rule out particle-object inversion, as when the object is inherently so heavy that it far outweighs the particle. Let us consider the following paraphrase pair.

(224) a. I gave *up teaching English.*
b. *I gave *teaching English up.*

The clausal object is much heavier than the particle here so that (224a) is in line with the weight-over-order principle while (224b) is not. This means that we should not allow particle-object inversion to apply to (224a) if we are to block the ungrammatical (224b).

Note that the derivation of the idiomatic *if you please* involves verb-object inversion. Let us consider the paraphrase pair below.

(225) a. Come this way, *if you please.*
b. (?)Come this way, *if it please(s) you.*

(225a) derives from something like (225b) via *it*-deletion and verb-object inversion. As we have already seen in 2.4., *it* is so light that it is often deleted as a consequence of the weight-over-length principle, as it is in the derivation of (225a) from (225b). As to the verb-object inversion, *you* is lighter than *please* here so that the two words trade places under the pressure of the weight-over-order principle. Suppose here that *you* in (225b) were replaced with, say, *Your Highness.* Then the object would be heavier than the verb so that the two would not swap places, which is also a consequence of the weight-over-order principle.

Note incidentally that in the derivation of (225a) from (225b) *you* may move over *please* to fill the vacuum left by the deletion of *it.* Thus the verb-object inversion here may be triggered by the deletion of the light-weight subject *it.* Recall that a similar phenomenon was observed in connection with (191) through (196) as well as (218).

The weight-over-order principle may also help explain the preposition-complement inversion illustrated by the paraphrase pair below.

(226) a. They did it *notwithstanding the bad weather.*
b. They did it *the bad weather notwithstanding.*

Notwithstanding is fairly heavy in that it comes from the clause *not* + *withstanding.* It is thus often inherently as heavy as its complement. When it receives information focus, *notwithstanding* is definitely heavier than its complement. Thus it often trades places with its complement in accordance with the weight-over-order principle, as in the derivation of (226b) from (226a), which is arguably triggered by the information focus on *notwithstanding.*

Similar in nature is the preposition-complement inversion exemplified by the following paraphrase pairs.

(227) a. *apart from* this weakness
b. this weakness *apart*
(228) a. *aside from* this weakness
b. this weakness *aside*

Apart and *aside* are inherently heavy of structural content in that they originate in prepositional phrases so that they are often as heavy as their complements. Besides, they may receive information focus so that they may gain extra semantic weight and thus become quite heavy. As a result, they may change places with their complements in accordance with the weight-over-order principle. This is apparently what happens in the derivation of the second expression from the first in either (227) or (228). Note that this derivation involves the deletion of *from,* which we will not discuss here as it is beyond the scope of the present study.

Suppose that *this weakness* were replaced with *discussing this weakness* in (227a) and (228a). Then the complement would become too heavy of structural content to allow the kind of preposition-complement inversion under discussion here. This again is a consequence of the weight-over-order principle, which incidentally can also help explain such inversion as is involved in extraposition of various types.

3.7. Binominals

The two terms of a binominal normally occur in ascending order of weight, i.e. in accordance with the weight-over-order principle. Let us take the following sentence pair for example.

(229) a. I can match the ballerina *step for dazzling step.*
b. *I can match the ballerina *dazzling step for step.*

Note that *dazzling step* is heavier than *step* so that the binominal phrase complies with the weight-over-order principle in (229a), but not in (229b), which is why (229a) is grammatical while (229b) is not.

Let us now consider the following binominals.

(230) a. rules and regulations
b. first and foremost
c. whys and wherefores

The first term of each binominal here is shorter and thus lighter than the second so that the two terms are ordered in accordance with the weight-over-order principle. On closer inspection, we notice that the two terms of each binominal here constitute an unmarked-marked pair such that the first term is umarked while the second term is marked. Note that the unmarked first term is more familiar to us and thus lighter of information content than the

marked second term. Thus from this perspective also, the two terms here are ordered in accordance with the weight-over-order principle.

Let us consider the following additional examples.

(231) a. right or wrong
b. pros and cons
c. hit or miss
d. make or break
e. up and down
f. in and out
g. with or without
h. plus or minus
i. yes or no
j. (through) thick and thin
k. (the) long and short (of the story)

The first term of each binominal here is positive and thus unmarked while the second term is negative and thus marked. As a result, the first term is lighter than the second so that the two terms are ordered in accordance with the weight-over-order principle.

The two terms of each binominal below also constitute an unmarked-marked pair.

(232) a. back and forth
b. to and fro
c. cup and saucer
d. bread and butter
e. gin and tonic

The first term of each binominal here is more unmarked than the second in that the former is lexically more familiar to us than the latter. Thus here again the first term precedes the second term in accordance with the weight-over-order principle.

Note that the two terms of a binominal numeral such as the old coordinate numeral below occur in ascending order of weight.

(233) *four* and *ten* (=fourteen)

Four is numerically smaller and thus lighter than *ten* so that the former precedes the latter in (233) in accordance with the weight-over-order principle.

The concept of binominal may be extended to include those binominals whose terms are in two different coordinate clauses, as in (234) below.

(234) a. *Finders,* keepers; *losers,* weepers.
b. You *win* some, and you *lose* some.

Note that *finders* and *win* here are unmarked terms while *losers* and *lose* are their respective marked counterparts. Note further that the unmarked terms here precede their respective marked counterparts in accordance with the weight-over-order principle.

The two terms of a binominal may be in two correlatively conjoined parts, as in the sentences below.

(235) a. Not only *John* but also *Jane* is to blame.
b. Neither *John* nor *Jane* is to blame.
c. Both *John* and *Jane* are to blame.

The speaker-writer takes *John* here as more given than *Jane* so that *John* is informationally lighter than *Jane*, that is, at least from the speaker-writer's point of view. (This appears to be especially the case in (235a).) Thus *John* precedes *Jane* in each sentence here in accordance with the weight-over-order principle.

The weight-over-order principle also operates on such binominals as those exemplified below, where correlative conjunction is less explicitly marked.

(236) a. He's not just *bright;* he's *brilliant.*
b. *He's not just *brilliant;* he's *bright.*
(237) a. It's not just *lovely;* it's *gorgeous.*
b. *It's not just *gorgeous;* it's *lovely.*

Bright and *lovely* are less intense in meaning, and thus semantically lighter, than *brilliant* and *gorgeous* respectively. Thus the first sentence in either pair here follows the weight-over-order principle while the second sentence does not, which is why the former is grammatical while the latter is not.

The exchange below illustrates an even less explicitly conjoined "binominal."

(238) A: Are you *sure*?
B: *Positive.*

Sure is less intense in meaning, and thus semantically lighter, than *positive* so that the former precedes the latter here in accordance with the weight-over-order principle.

It is interesting that *who* and *what* are also normally conjoined in ascending order of weight, as can be seen from (239) below.

(239) a. I was wondering *who* and *what* he was.
b. ?I was wondering *what* and *who* he was.

Note that *who,* which is of human reference, is informationally lighter than *what,* which is of non-human reference. Thus (239a) is more in line with the weight-over-order principle than (239b), which is why the former is more natural than the latter.[7]

Binominals of the type exemplified below also normally follow the weight-over-order principle.

(240) a. Anglo-American
b. *Americo-English
(241) a. Sino-Korean
b. *Koreano-Chinese

The hyphen in these binominals is equivalent to the coordinator in explicit binominals. Note that the first term is shorter and thus lighter than the second in the first member of either pair here while it is the other way around in the second member of the same pair. Thus the first member of either pair here follows the weight-over-order principle while the second member does not, which is why the former is grammatical while the latter is not.

Needless to say, other factors than weight may sometimes intervene, as in the following examples.

(242) a. Judaeo-Christian
b. *Christo-Judaean

[7]That the human *who* is lighter than the non-human *what* also helps explain why "*Who* has *what*?" is perfectly natural while "*What* does *who* have?" is definitely odd.

(243) a. Anglo-French
b. *Franco-English

Note that *Judaeo-* and *Anglo-* are more polysyllabic and thus phonologically heavier than *Christian* and *French* respectively. Thus *Judaeo-Christian* and *Anglo-French* apparently violate the weight-over-order principle so that they should be ungrammatical or at least odd. However, they are both perfectly grammatical and natural. The reason for this is that chronological and ethnocentric factors intervene here and nullify the weight-over-order principle. Judaism is chronologically prior to Christianity so that *Judaeo-Christian* follows this relation of chronological precedence. *Anglo-* is, for the English people at least, ethnocentrically prior to *French* so that *Anglo-French* follows this relation of ethnocentric precedence.

It may be the case that a precedence relation is such that something that is prior in that relation is more given than something that is not. If this is indeed the case, then we may argue that examples such as *Judaeo-Christian* and *Anglo-French* also follow the weight-over-order principle so that they do not constitute genuine counterexamples to the principle.

The expressions *and/or* and *either/or* are also implicit binominals. Note that the two terms of these binominals are also ordered in accordance with the weight-over-order principle. For *and* and *either* are positive and thus unmarked while *or* is negative and thus marked, so that *and* and *either* are lighter than *or*.

Note at this point that most reduplications are implicit binominals and that they also follow the weight-over-order principle. Let us take the following reduplications for example.

(244) a. zigzag
b. drip-drop
c. heehaw

Note that the vowel is higher and thus lighter in the first syllable than in the second in each reduplication here so that the first syllable is lighter than the second. Thus the two terms of each binominal here, i.e. the two syllables comprising each reduplication here, are ordered in accordance with the weight-over-order principle.

Explicit reduplicative binominals such as the following also comply with the weight-over-order principle.

(245) a. tit for tat
b. spic and span
c. dribs and drabs
d. hem and haw
e. quid pro quo

The vowel is higher and thus lighter in the first term of each reduplicative binominal here than in the second. Thus the first term is lighter than the second so that the two terms here are ordered in accordance with the weight-over-order principle.

It is interesting that the weight-over-order principle can help explain why *mickle* precedes *muckle* in the following proverb.

(246) Many a *mickle* makes a *muckle*.

Mickle is phonologically lighter than its variant *muckle* in that the vowel is lighter in the former than in the latter. Thus *mickle* precedes *muckle* in the proverb here in accordance with the weight-over-order principle.

Note that the (consonantal) onset is normally heavier in the second element than in the first in rhyming reduplications such as those below.

(247) a. *s*uper-*d*uper
b. *h*anky-*p*anky
c. *r*ag*t*ag
d. *ϕ* even-*st*even
e. *ϕ* okey-*d*okey
f. *w*ine and *d*ine
g. *wh*eel and *d*eal

Note that the two elements of each reduplication here occur in ascending order of phonological weight, that is, in compliance with the weight-over-order principle.

Given our comments on (247) above, we can easily explain why *hustle* precedes *bustle* in the italicized idiomatic phrase below.

(248) I don't like the *hustle and bustle* of the city.

Note that *hustle* is lighter than *bustle* in that /h/ is phonologically lighter than /b/. Thus *hustle* precedes *bustle* in accordance with the weight-over-

order principle.

The weight-over-order principle sometimes operates on binominals on the basis of extralinguistically defined weight, as in the examples below.

(249) a. now and then (=occasionally)
b. *then and now (=ditto)
(250) a. here and there (=at various times or places)
b. *there and here (=ditto)
(251) a. this and that (=various things)
b. *that and this (=ditto)

Now, here, and *this* are referentially closer to the immediate context of speech than *then, there,* and *that* respectively so that the former are informationally, and thus phonologically, lighter than the latter. Thus the first member of each pair above follows the weight-over-order principle while the second does not, which is why the former is grammatical while the latter is not. Incidentally, the starred forms here could become grammatical if other factors intervened. For example, (249b) would be perfectly grammatical if two points in time, i.e. then and now, are listed in chronological order. Note here that the first member of each pair above can only have an idiomatic interpretation while the second member, if and when unstarred, can only have a non-idiomatic interpretation.

Note that the two time expressions in each proverb below occur in ascending order of weight.

(252) a. Here *today* and gone *tomorrow.*
b. An egg *today* is better than a hen *tomorrow.*
c. *Now* is now and *then* was then.

Today and *now* are referentially closer to the immediate context of speech and thus informationally lighter than *tomorrow* and *then* respectively. Thus *today* and *now* precede *tomorrow* and *then* respectively here in accordance with the weight-over-order principle.

The two place expressions in each proverb below also occur in ascending order of weight.

(253) a. A bird *in the hand* is worth two *in the bush.*
b. A feather *in the hand* is better than a bird *in the air.*

In the hand is referentially closer to the immediate context of speech and thus informationally lighter than either *in the bush* or *in the air*. Thus *in the hand* precedes *in the bush* or *in the air* here in compliance with the weight-over-order principle.

Incidentally, we need a proviso here to the effect that the negation of the predicate in sentences such as those of (253) may override the application of the weight-over-order principle. For (253a), for one, may be paraphrased as "Two birds *in the bush* are *not* worth one (bird) *in the hand,*" in which the order of the adverbials is exactly the opposite of that in the source sentence, i.e. (253a).

4. Weight and Rhythm

4.0. The Weight-Over-Rhythm Principle

The rhythm of English, which is definable in terms of weight, normally follows the weight-over-rhythm principle given below.

> HEAVY ELEMENTS ALTERNATE WITH LIGHT ELEMENTS IN SUCH A WAY THAT ELEMENTS OF COMPARABLE WEIGHT, ESPECIALLY THOSE OF HEAVY WEIGHT, TEND NOT TO BE PROXIMATELY REPEATED.[8]

This principle has the effect of keeping elements of similar weight, especially heavy ones, from bunching together. Clusters of similar elements, including elements of similar weight, may pose something of a problem for language processing (Park 1977a, 1977b, 1982, 1983, and 1984). This would especially be the case with clusters of heavy elements. Thus it would appear that the weight-over-rhythm principle may serve to ease the burden on the language user by preventing the formation of such problematic clusters.

4.1. Stress Rhythm

As already noted above, the weight-over-rhythm principle discourages clusters of tonic syllables from forming.[9] It often uses function words to do this, as can be seen from (254) below.

[8]For a detailed discussion of the constraint on proximate repetition of similar elements, see Park (1977b, 1982, 1983 and 1984).

[9]A syllable is said to be tonic if it receives non-zero stress.

(254) a. *Mȳ* nâme *ĭs* Jóhn.
b. *Shĕ* rêad *hĭs* bóok.

Note that tonic content words alternate quite regularly with atonic function words here. Although the alternation is not always this regular, it is by and large regular so that it bears out the weight-over-rhythm principle.

Examples such as the following also attest to the role that function words play in keeping tonic syllables from clustering together.

(255) a. nîce *ănd* wárm (=warm in a nice way)
b. gȯȯd *ănd* wárm (=warm in a good way)

Note that *nice* and *good* intensify *warm* here and are thus adverbial in function so that *and* may be regarded as an adverbial "suffix" of sorts. Thus *and* here may be viewed as an atonic "suffixal" buffer between two tonic content words.

Bona fide suffixes also serve as a buffer between two tonic elements. Let us take the following expressions for example.

(256) a. dêad*lȳ* blóws
b. cômĭc stríps
c. ea̍rthĕn póts

Each suffix here comes between two tonic elements, i.e. between the stem to which it is suffixed and the noun that follows. Thus it serves to block a juxtaposition of two tonic elements, which would otherwise seriously violate the weight-over-rhythm principle.

The importance of the suffix as a buffer is clearly demonstrated by paraphrase pairs such as the following.

(257) a. mȳ wêddĕd wífe
b. *mȳ wêd wífe
(258) a. ă lîghtĕd cándl̆e
b. *?ă lît cándl̆e

Both *wed* and *light* have two alternative past participial forms, one suffixed and one unsuffixed. Note that only the suffixed form is acceptable when the participle is used (by itself) as an attributive adjective. This is because the suffixed form, but not the unsuffixed form, provides a buffer between the

two tonic elements, i.e. the stem of the adjective and the noun that follows. Note that the second member of either pair above violates the weight-over-rhythm principle while the first member of the same pair does not.

The weight-over-rhythm principle also provides a principled explanation as to why *-teen,* but not *-ty,* is tonic in numeral pairs such as the following.

(259) a. fífteén
b. fífty̆

Note that numerals in *-ty* are routinely "suffixed" with a basic numeral, i.e. any numeral from *one* through *nine,* while numerals in *-teen* are never so suffixed. Suppose that *-ty* were allowed to get strong stress. Then a cluster of three tonic syllables would form in such numerals as *fifty-five* and *fifty-six*. For basic numerals always get strong stress on the first syllable. Thus we assign weak stress, i.e. zero stress, to *-ty* so as to keep numerals such as *fifty-five* and *fifty-six* from seriously violating the weight-over-rhythm principle. Note that *-teen* is not under similar pressure to weaken its stress.

The weight-over-rhythm principle can also help explain why the first member of each pair below is favored over the second.

(260) a. nóoks ǎnd cránnĭes
b. ?cránnĭes ǎnd nóoks
(261) a. ládĭes ǎnd géntl̆emĕn
b. ?géntl̆emĕn ǎnd ládĭes

Note that the first member of either pair here involves less of a cluster of atonic syllables than does the second. Thus the first member is more in line with the weight-over-rhythm principle than the second, which is one reason why the former is preferred to the latter.

Incidentally, stress rhythm is a surface manifestation of semantic rhythm in that the strength of stress is a function of the weight of meaning. Note in this connection that a higher-level semantic rhythm also manifests itself, as in the regular alternation between subject and predicate, which normally bear old and new information respectively.

Stress rhythm operates not just across word boundaries, as in the examples we have so far considered. It also operates intralexically, as in the following examples.

(262) a. ĕxàmĭnátiŏn
b. ècŏnómĭc
c. Hémĭngwày

Tonic syllables alternate quite regularly with atonic syllables in the words here. Although the alternation is not always this regular, it is by and large quite regular in much of the English lexicon so that intralexical stress also bears out the weight-over-rhythm principle.

Note at this point that stress rhythm also figures in word-formation processes such as that exemplified below.

(263) a. ădmíre → àdmĭrátiŏn
b. ĕxpóse → èxpŏsítiŏn
c. cŏndémn → còndĕmnátiŏn

The stress pattern changes from verb to noun in each pair here in such a way that neither strong nor weak stress is proximately repeated in the resulting noun. Note that the rhythmic distribution of stress associated with the word-formation process in question here is in line with the weight-over-rhythm principle.

Note that the infix *-bloody-* is so positioned in the word below as to guarantee a rhythmic distribution of stress.

(264) ábsŏ*blòodỹ*lútelỹ

Positioning *-bloody-* between *abso-* and *-lutely* here guarantees two things. Firstly, the original stress pattern remains intact in both the matrix word and the infix. Secondly, neither strong nor weak stress is proximately repeated in the resulting word so that stress gets rhythmically distributed in line with the weight-over-rhythm principle.

The linking *-o-* and *-a-* are also so positioned in the words below as to guarantee a rhythmic distribution of stress in the resulting words.

(265) a. Brăzílŏphìle
b. wòrkăhólĭc

Were it not for the atonic *-o-* and *-a-* here, we would end up with a cluster of two tonic syllables in either word above. Standing between the two tonic syllables here, the linking *-o-* and *-a-* serve to make the stress in the result-

ing words better comply with the weight-over-rhythm principle.

Note that the change in stress pattern illustrated below is also rhythmically motivated.

(266) évĭdĕnt → évĭdèntlȳ/èvĭdéntlȳ/évĭdĕntlȳ

Note that the third syllable may be tonic in the adverb here, but not in the adjective. Note further that, were it not for the tonicity of the third syllable, *evidently* would end in a cluster of three atonic syllables. Thus the optional strong stress on the third syllable of *evidently* is apparently designed to make the stress here more rhythmic than would otherwise be the case. Similar in motivation may be the primary stress (in American English) on the antepenultimate syllables of *primarily, militarily,* etc., as opposed to the non-primary stress on the penultimate syllables of *primary, military,* etc.

4.2. Stress-Cluster Simplification

Both strong- and weak-stress clusters, especially the former, violate the weight-over-rhythm principle and are thus maximally constrained. However, such clusters do occur, especially weak-stress clusters. When they do, they tend to be simplified in one way or another so as to alleviate their violation of the weight-over-rhythm principle.

Strong-stress clusters are very rare in English. Those that do occur are almost always broken up by means of pauses inserted between the tonic syllables in question. Since a pause is equivalent to an atonic syllable, so to speak, strong-stress clusters are even more of a rarity than one might suspect. At any rate, the use of pauses to break up strong-stress clusters is a consequence of the weight-over-rhythm principle.

Weak-stress clusters are tolerated to a far greater extent than strong-stress clusters. In fact, weak-stress clusters are fairly common. However, there are indications of general resistance to weak-stress clusters, especially to those of more than two atonic syllables. For example, clusters of three atonic syllables are almost always simplified to clusters of just two such syllables, as can be seen from the following examples.

(267) a. economically /èkə̆nə́mĭk(ə̆)lĭy/
b. comfortable /kʌ́mf(ə̆)tə̆bə̆l/
c. vegetable /védʒ(ə̆)tə̆bə̆l/

Note that the parenthesized vowels here are more often than not deleted. When they are not deleted, each word here ends in three consecutive atonic syllables so that its stress pattern is in slight violation of the weight-over-rhythm principle. Thus the deletion of the parenthesized vowels here serves to render the stress pattern more compliant with the weight-over-rhythm principle than would otherwise be the case.

It is interesting that the second token of the orthographic *e* in *vegetable* is almost always mute, which apparently speaks for the power of the weight-over-rhythm principle. Note in this connection that the muteness of the orthographic *i* in *parliament* could be viewed from a similar perspective.

The desire to simplify weak-stress clusters is apparently at the root of the alternation between *-ery* asnd *-ry* illustrated by the two sets of examples below.

(268) a. fish + -ery → físhĕrȳ
b. cook + -ery → coókĕrȳ
(269) a. gallant + -ry → gállăntrȳ
b. yeoman + -ry → yeómănrȳ

Note that *-ry* is a reduced version of *-ery* and that the former is used with disyllabic stems while the latter is used with monosyllabic stems. Suppose that we used *-ery* with the disyllabic stems in (269). Then the words here would end in three consecutive atonic syllables. Thus the use of *-ry* in (269) helps block the formation of a cluster of three atonic syllables, thereby rendering the resulting words more observant of the weight-over-rhythm principle than would otherwise be the case.

There is some resistance to clusters of just two atonic syllables also so that they often get simplified, that is, under the pressure of the weight-over-rhythm principle. When they get simplified, they do so by dropping the vowel in the first of the two atonic syllables in question. Let us take the following couplets for example.

(270) a. phántăsȳ → fáncȳ
b. cóurtĕsȳ → cúrtsȳ
c. pérĭloŭs → párloŭs

Historically, the first word gave rise to the second in each pair here. Note that the first word ends in two consecutive atonic syllables and is thus in

slight violation of the weight-over-rhythm principle. On the other hand, the second word ends in just one atonic syllable so that it is no longer in violation of the principle. Thus the derivation exemplified above involves a weak-stress cluster simplification apparently motivated at least in part by the weight-over-rhythm principle. Note incidentally that *parl-* in *parlous* is short for, and thus carries the weight of, *peril-* in *perilous*. Thus the nucleus is heavier and longer in the first syllable of *parlous* than in that of *perilous*, that is, in *par-* than in *per-*. This is in line with the weight-over-length principle.

Note in this connection that *medicine* alternates between /médəsən/ and /médsən/. We may say that /médsən/ derives from /médəsən/ in the same way that the second word in each pair of (270) derives from the first, that is, under the pressure of the weight-over-rhythm principle. Incidentally, *median* must also have given rise to *mean* under the pressure of the weight-over-rhythm principle. Note that the pre-atonic *d* apparently got so light here that it ended up getting deleted along with the *i* that follows (See 2.4.).

Similarly explainable is the historical development of the following words from trisyllables to disyllables.

(271) a. Leicester
b. Worcester
c. Gloucester

Each word here originally ended in two consecutive atonic syllables, the first of which lost its vowel under the pressure of the weight-over-rhythm principle. With the loss of this vowel the tonic first vowel now came to precede a consonant cluster, which explains its laxing in *Leicester* and *Gloucester* (See 4.3.). Note also that the orthographic *c* (=/s/) merged with the following *s* (=/s/) so that only one /s/ is pronounced today.

The desire to simplify weak-stress clusters has apparently contributed to the historical development of the following words from trisyllables to disyllables.

(272) a. mission
b. vision
c. partial

All these words orginally terminated with two consecutive atonic syllables so that the orthographic *i* in the immediate post-tonic syllable was not mute.

Partly under the pressure of the weight-over-rhythm principle, however, this *i* weakened to /y/, which then palatalized the immediately preceding consonant and then zeroed out.[10]

The following words underwent a similar historical change from trisyllabic to disyllabic words.

(273) a. onion
b. million

These words originally ended in two consecutive atonic syllables. The orthographic *i* in the first of the two atonic syllables weakened to /y/ partly under the pressure of the weight-over-rhythm principle. This /y/ has not, however, completely palatalized the immediately preceding consonant and thus has not zeroed out. Note that /y/ is not nucleic so that the words in (273) end in just one atonic syllable in present-day English.

The weight-over-rhythm principle may also be instrumental in explaining why the orthgraphic *a* in the words below is, more often than not, mute.

(274) a. ridd*a*nce
b. pitt*a*nce
c. med*a*l
d. met*a*l

Note that /n/ and /l/ here are so vocalic that they may arguably be taken to be extremely weak atonic vowels. Under this assumption, we can say that each word here ends in two consecutive atonic syllables and that the first of these two syllables has lost its vowel under the pressure of the weight-over-rhythm principle. The loss of the vowel here is thus essentially identical to that in (270) through (273). Needless to say, the silencing of the orthographic *e* and *o* in *written, ridden, idol,* etc. can also be explained in exactly the same way.

The behavior of the italicized *i* and *e* in the word pairs below are also apparently motivated at least in part by rhythmic considerations.

(275) a. sóc*i*ăl vs. sòc*ĭ*ólŏgy̆
b. párt*i*ăl vs. pàrt*ĭ*álĭty̆
c. óc*e*ăn vs. òc*ĕ*ánĭc

[10]As to why this /y/ zeroed out, see Park (1983).

The italicized *i* and *e* here are nucleic in the second member of each pair here while they are not in the first member of the same pair. Were this not the case, then the first member of each pair here would end in a cluster of two atonic syllables while the second member of the same pair would be encumbered with a cluster of two tonic syllables. Thus whether or not the italicized vowel letter is realized as a nucleus here is at least in part a consequence of the weight-over-rhythm principle. A similar account applies to the fact that the second orthographic *i* is often nucleic in *familiarity*, but never in *familiar*.

The weight-over-rhythm principle also throws light on the interesting behavior of auxiliary contractions, as illustrated by the following paraphrase pairs.

(276) a. Hĕ hăs dóne ĭt.
b. Hĕ's dóne ĭt.
(277) a. Jôhn hăs dóne ĭt.
b. Jôhn's dóne ĭt.

Note that (276a) involves a cluster of two atonic syllables while (276b) does not. Thus (276b) is more in line with the weight-over-rhythm principle than (276a) so that the former is more natural than the latter. On the other hand, (277b) involves a cluster of two tonic syllables while (277a) does not. Thus (277a) is more in line with the weight-over-rhythm principle than (277b) so that the former is more natural than the latter. Note here that the weight-over-rhythm principle is instrumental in explaining why an auxiliary contraction is normally more natural when suffixed to a pronoun than when suffixed to a noun.

4.3. Segmental Rhythm

The weight-over-rhythm principle operates on sequences of segmental sounds in such a way that heavy segments or clusters of segments regularly alternate with light ones. Thus a heavy stem vowel is generally followed by a light consonantal coda while a light stem vowel is followed by a heavy consonantal coda. Let us consider the following examples.

(278) a. bat /bæt/ (= [bæt]) vs. bad /bæd/ (= [bæ·d])
b. back /bæk/ (= [bæk]) vs. bag /bæg/ (= [bæ·g])

Note that one and the same vowel is phonetically longer and thus heavier before a voiced consonant than before a voiceless one here. Note further that a voiceless consonant is heavier than its voiced counterpart. Thus one and the same vowel is heavier here before a lighter consonantal coda than before a heavier one, which is in line with the weight-over-rhythm principle.

Segmental rhythm manifests itself even more clearly in the following pairs of words.

(279) a. deep /d*iyp*/ vs. depth /d*epθ*/
b. heal /h*iyl*/ vs. health /h*elθ*/
c. scribe /skr*ayb*/ vs. script /skr*ipt*/

The vocalic nucleus is more complex and thus heavier in the first word of each pair here than in the second. On the other hand, the consonantal coda is more complex and thus heavier in the second word of each pair than in the first. Thus the weight of the nucleus here is in inverse proportion to that of the coda, which is in line with the weight-over-rhythm principle. Note incidentally that the vocalic laxing illustrated by (279) is a consequence of the weight-over-rhythm principle.

Note that the explanation of the following examples hinges on the premise that a voiceless consonant is heavier than its voiced counterpart.

(280) a. dream: dreamed /dr*iymd*/ vs. dreamt /dr*emt*/
b. cleave: cleaved /kl*iyvd*/ vs. cleft /kl*eft*/

In either word set here, the two alternative past (participial) forms terminate with a consonant cluster. This cluster is lighter in the first form than in the second in that it is voiced in the former while it is at least in part voiceless in the latter. Note here that the nucleus is heavier in the first form than in the second. Thus here again the weight of the nucleus is in inverse proportion to that of the coda, which accords with the weight-over-rhythm principle.

The explanation of the following examples also hinges on the premise that a voiceless consonant is heavier than its voiced counterpart.

(281) a. h*ind*, m*ind*, b*ind* /*aynd*/
b. h*int*, m*int*, fl*int* /*int*/
(282) a. ch*ild*, m*ild*, w*ild* /*ayld*/
b. h*ilt*, k*ilt*, t*ilt* /*ilt*/

Note that the heavier nucleus /ay/ is followed by the lighter coda /nd/ or /ld/ while the lighter nucleus /i/ is followed by the heavier coda /nt/ or /lt/. This is perfectly in line with the weight-over-rhythm principle. Incidentally, it may be in order here to note that there are some exceptions such as *pint* /paynt/ and *wind* /wind/ here.

The weight-over-rhythm principle affords a rare insight into the voicing phenomenon exemplified by the singular-plural pairs below.

(283) a. mouth /ma*wθ*/ vs. mouths /ma*wðz*/
b. oath /*owθ*/ vs. oaths /*owðz*/ or /*owθs*/
c. sheath /*ʃiyθ*/ vs. sheaths /*ʃiyðz*/ or /*ʃiyθs*/

Note that the nucleus in every word here is diphthongal and thus heavy. Note further that this heavy nucleus is followed by a single consonant in the singular and by a cluster of two consonants in the plural. Thus the plural form here is slightly out of line with the weight-over-rhythm principle.[11]

Note at this point that the final cluster in the plural form here gets voiced most of the time. Note further that voiced consonants are lighter than their voiceless counterparts so that voicing a consonant has the effect of making it lighter. Thus the voicing of the plural-final cluster in (283) serves to lighten the cluster so that the plural form as a whole may be more in line with the weight-over-rhythm principle than would otherwise be the case.

Note also that the plural-final cluster in (283) is more likely to get voiced when the nucleus is more clearly diphthongal and thus heavier than when it is less so. Thus the voicing here is obligatory for *mouths* whereas it is optional for *oaths* and *sheaths*. This is apparently a consequence of the weight-over-rhythm principle, which requires that the nucleus be weightwise in inverse relation to the coda.

Similarly explainable is the voicing of the plural-final cluster exemplified by the following singular-plural pairs.

(284) a. kinfe /na*yf*/ vs. knives /na*yvz*/
b. loaf /l*owf*/ vs. loaves /l*owvz*/
c. leaf /l*iyf*/ vs. leaves /l*iyvz*/

Note at this point that the plural-final cluster does not normally get

[11]The weight-over-rhythm principle is violated only slightly here because the plural-final consonant cluster in question is not that heavy as consonant clusters go. This is because the cluster comprises only fricatives, which are relatively light consonants.

voiced in the examples below.

(285) a. roo*f* vs. roo*fs*
b. oa*f* vs. oa*fs*

The nucleus here is rather weakly diphthongal and thus relatively light, which is apparently why the plural-final coda is left unvoiced and thus heavier than it would otherwise be. This must also be a consequence of the weight-over-rhythm principle. However, the mystery still remains as to why the same cluster gets voiced in similar environments, as in (284b) and (284c).

Note in this connection that *hoof* has two alternative plural forms, as shown below.

(286) hoofs /h*ufs*/ or hooves /h*uwvz*/

Note that /uw/ is heavier than /u/ while /fs/ is heavier than /vz/ so that the heavier nucleus /uw/ pairs with the lighter coda /vz/ here while the lighter nucleus /u/ pairs with the heavier coda /fs/. Note that this pairing complies with the weight-over-rhythm principle.

The voicing of the plural-final cluster exemplified below is essentially identical in nature to that illustrated by (284) and (286).

(287) a. self vs. selves
b. wolf vs. wolves
c. scarf vs. scarves or scarfs
d. turf vs. turves or turfs

Note that the postvocalic /l/ is very much like /w/ while the postvocalic /r/ is very much like /ə/. Thus the vocalic nulceus in every pair of (287) is diphthongal, so to speak, and is thus heavy enough to trigger the lightening via voicing of the plural-final cluster here. Note at this point that the nucleus is more clearly diphthongal and thus heavier in the *l* words here than in the *r* words. As a result, the nucleus exerts more pressure on the plural-final cluster to voice and lighten in the *l* words here than in the *r* words. This is why voicing is obligatory in the *l* words here while it is optional in the *r* words. Needless to say, this is all in line with the weight-over-rhythm principle.

Note in this connection that the plural-final cluster does not get voiced in words such as the following.

(288) a. belie*f* vs. belie*fs*
b. sa*fe* vs. sa*fes*

There seem to be two reasons for the failure of the plural-final cluster here to get voiced. One is that the words in question came into English well after the *f-v* voicing had run its course. The other reason is that the voicing here was probably blocked by such already existing verb forms as *believes* and *saves*.

It is interesting that genitive-final consonant clusters are often quite similar in form to plural-final consonant clusters and yet do not undergo lightening via voicing. Let us take the following pairs for example.

(289) a. wol*f* vs. wol*f's*
b. Sou*th* vs. Sou*th's*

Note that the genitive *-'s* is less tightly bound to the stem than is the plural *-s* so that the stem-final *f* or *th* forms a less tight and thus lighter cluster with *-'s* than with *-s*. Thus the weight-over-rhythm principle exerts less pressure on the genitive-final cluster to lighten via voicing than it does on the plural-final cluster. This is apparently why the genitive-final cluster here does not get voiced while a similar plural-final cluster normally does.

It is interesting that cluster tightness is also a factor in the vowel alternation exemplified by (290) below.

(290) a. cleans /kl*iynz*/ (as in "He *cleans* his room every day.")
b. cleanse /kl*enz*/

The word-final suffix in *cleans* is far more productive than that in *cleanse* with the result that the suffix is perceived as such far more readily in *cleans* than in *cleanse*. Thus the suffix is more tightly bound to the stem in *cleanse* than in *cleans* so that /nz/ is a tighter and thus heavier cluster in the former than in the latter. Thus the heavier /nz/ follows the lighter nucleus /e/ here while the lighter /nz/ follows the heavier nucleus /iy/. Note that this accords with the weight-over-rhythm principle.

The stem-vowel alternation illustrated by the following word pair is also keyed to the tightness of the consonant cluster that follows the stem vowel.

(291) a. cleanly /kl*iynl*iy/ (adverb)
b. cleanly /kl*enl*iy/ (adjective)

Note here that *-ly* is productive as an adverbial suffix while it is unproductive as an adjectival suffix. Thus /nl/ is more tightly bound and thus heavier in (291b) than in (291a). As a result, the heavier /nl/ follows the lighter nucleus /e/ here while the lighter /nl/ follows the heavier nucleus /iy/, which again is in line with the weight-over-rhythm principle.

Note that the vowel alternation illustrated below is keyed to the manner of articulation of the word-final consonant.

(292) a. mean: means /m*iynz*/ vs. meant /m*ent*/
b. keep: keeps /k*iyps*/ vs. kept /k*ept*/
c. feel: feels /f*iylz*/ vs. felt /f*elt*/

Note that fricatives are much lighter than stops so that /nz/, /ps/, and /lz/ are much lighter than /nt/, /pt/, and /lt/ respectively. Thus the heavier nucleus /iy/ pairs with the lighter codas here while the lighter nucleus /e/ pairs with the heavier codas, which is in line with the weight-over-rhythm principle.

Note in this connection that consonant clusters ending in a fricative do not always follow a light nucleus. Let us take the following for example.

(293) heal: heals /h*iylz*/ vs. health /h*elθ*/

Note that *-s* here is a productive suffix while *-th* is an unproductive one so that /lz/ is a less tightly bound, and thus a less heavy, cluster than /lθ/. Thus /lθ/ exerts more pressure on the nucleus to lighten via laxing here than does /lz/, which is probably why the nucleus is heavier before /lz/ than before /lθ/ in (293). It may also be the case that /θ/ is inherently heavier than /z/ so that /lθ/ is an inherently heavier cluster than /lz/. Either way, the vowel alternation here is tied to the weight of the word-final consonant cluster in such a way that it is in line with the weight-over-rhythm principle.

The two variants of the tonic vowel in each pair below can also be explained in terms of the weight-over-rhythm principle.

(294) a. sane /s*ey*n/ vs. sanity /s*æ*nitiy/
b. serene /sir*iy*n/ vs. serenity /sir*e*nitiy/

Note that we can extend the concept of coda to include atonic (poly)syllabic suffixes such as *-ity* here. Under this extended concept of coda, the coda is the monoconsonantal /n/ in the first member of either pair here while it is

this /n/ plus *-ity* in the second member of the same pair. Since /ey/ and /iy/ are heavier than /æ/ and /e/ respectively, we can say that the heavier coda *-nity* follows the lighter nucleus /æ/ or /e/ while the lighter coda /n/ follows the heavier nucleus /ey/ or /iy/. Note that this is all in line with the weight-over-rhythm principle. Incidentally, we can now readily explain the /iy/~/e/ alternation between *please* and *pleasure/pleasant*. Note in this connection that *pleasing* differs from either *pleasure* or *pleasant* in that its stem vowel is /iy/, not /e/. This is because *-ing* is more productive and thus less tightly bound to the stem here than is either *-ure* or *-ant* so that *-ing* is a lighter suffix than either *-ure* or *-ant*. Needless to say, this is also in line with the weight-over-rhythm principle.

Note that the vowel alternation in the tonic syllables of the following words is also keyed to the weight of the coda.

(295) a. private /prá*y*vĭt/
b. privative /prívə̆tĭv/
c. privacy /prá*y*və̆sĭy/ or /prĭvə̆sĭy/

The tonic syllable here has a monosyllabic coda in the first word while it has a disyllabic coda in the second and third words. Thus the coda is heavier in the second and third words than in the first. Note also that the coda is heavier in the second word than in the third in that it contains a stop and two fricatives in the former while it contains only two fricatives in the latter. Note here that the lightest coda *-vate* follows the heavier nucleus /ay/ while the heaviest coda *-vative* follows the lighter nucleus /i/. Note also that *-vacy,* the coda that is in between in weight, follows either the heavier nucleus or the lighter one. This pairing is all in line with the weight-over-rhythm principle.

The extended concept of coda also throws light on the vowel alternation illustrated by the following pair of words.

(296) a. graph /gr*a:f*/ or /gr*æf*/
b. graphic /gr*æfik*/

The stem vowel is often /ɑ:/ in *graph,* especially in British English, while it is always /æ/ in *graphic*. Thus the heavier stem vowel /ɑ:/ here is followed by the lighter coda /f/ while the lighter stem vowel /æ/ is followed by the heavier coda /fik/, which is in line with the weight-over-rhythm principle.

Note in this connection that the suffix *-y* is not quite as heavy as the suffix

-ic. Let us compare (297) below with (296) with respect to stem-vowel alternation.

(297) a. class /kl*a:s*/ or /kl*æs*/
b. classy /kl*a:siy* / or /kl*æsiy*/

The stem vowel here tends to remain constant for any given speaker of English so that the suffix *-y*, unlike *-ic* in (296), does not add enough weight to the coda to lighten the preceding stem vowel. This is evidently because *-y* is not quite as heavy as *-ic*. Needless to say, this is also in line with the weight-over-rhythm principle.

It may be in order here to point out that the American English stem vowel /æ/ often corresponds in British English to /ɑ:/ before fricatives and to /æ/ elsewhere. Thus British English tends to use the heavier /ɑ:/ before lighter codas here while it tends to use the lighter /æ/ before heavier codas, which is in line with the weight-over-rhythm principle.

Segmental rhythm is sometimes achieved by means of epenthesis. Let us consider the followig examples.

(298) a. confessed /kənf*ést*/ vs. confessedly /kənf*ésidl*iy/
b. ashamed /ə∫*éymd*/ vs. ashamedness /ə∫*éymidn*is/
(299) a. accustomed /əkʌst*ə̆md*/ vs. accustomedness /əkʌst*ə̆mdn*is/
b. goodnatured /gudneyt∫*ə̆rd*/ vs. goodnaturedly /gudneyt∫*ə̆rdl*iy/

Note that the nucleus of the adjective-final syllable here is tonic in (298) while it is atonic in (299) so that it is heavier in the former than in the latter. Note also that, when followed by either *-ly* or *-ness,* the adjective-final consonant cluster is broken up by an epenthetic vowel in (298), but not in (299). Suppose that the cluster were not thus broken up in (298). Then we would end up with a heavy nucleus followed by a heavy cluster of three consonants in *confessedly* and *ashamedness*. This would violate the weight-over-rhythm principle rather seriously. On the other hand, the three-member consonant cluster in either *accustomedness* or *goodnaturedly* does not pose that much of a problem. For it follows a lighter nucleus so that it does not violate the weight-over-rhythm principle in any serious way. Thus the use of an epenthetic vowel in (298), as opposed to its non-use in (299), is arguably a consequence of the weight-over-rhythm principle.

Segmental rhythm may also be achieved by means of elision, as can be seen from the following examples.

(300) a. há*nds*ome /nds/ → /ns/
b. exá*ctl*y /ktl/ → /k(t)l/
c. á*sked* /skt/ → /s(k)t/

Suppose that the cluster-medial stop were not elided from the words here. Then we would end up with a tonic nucleus followed by a three-member consonant cluster. Since both the nucleus and the cluster, especially the latter, are quite heavy, their juxtaposition here would violate the weight-over-rhythm principle rather seriously. Thus the elision in (300), whether obligatory or optional, is arguably a consequence of the weight-over-rhythm principle.

The conjugation of some strong verbs such as the following may also involve a similarly motivated elision.

(301) a. think vs. thought (=*thinked /θiŋkt/)
b. drink vs. drank (=*drinked /driŋkt/)

If the verbs here were weakly conjugated, then their past (participial) forms would end in a cluster of at least three consonants, i.e. /ŋkt/. Since the nucleus is tonic and thus heavy, this would violate the weight-over-rhythm principle rather seriously. Thus part of the word-final cluster may have been elided here under the pressure of the weight-over-rhythm principle.

Note at this point that in (298) through (301) we have treated as heavy even lax vowels as long as they are tonic. These tonic lax vowels may not be quite as heavy as tonic tense vowels, but they are still heavy enough to pressure the following three-member consonant clusters to break up and simplify. Thus we may say that a tonic lax vowel can be treated as heavy for our purposes here when it is followed by a heavy cluster of three or more consonants. We are suggesting here that the weight of a nucleus is often relative to that of the coda that follows.

Note that the conjugation of the verbs below may be thought of as involving an elision similar to that discussed in connection with (301).

(302) a. sing vs. sang/sung (=*singed /siŋd/)
b. ring vs. rang/rung (=*ringed /riŋd/)

The starred forms here may be thought of as ending in three consecutive consonants. For there is reason to believe that /ŋ/ is underlyingly a sequence of two consonants, i.e. /n/ and /g/. For one thing, /ŋ/ historically

originates in /ng/. For another, /ŋ/ never follows a diphthongal nucleus intramorphemically. Given the weight-over-rhythm principle, we may interpret this as meaning that /ŋ/ is underlyingly polysegmental. On the strength of this much evidence, we may argue that the verbs in (302) are strongly conjugated partly under the pressure of the weight-over-rhythm principle. Note that the past (participial) forms in (301) arguably end in four, rather than three, consecutive consonants in that /ŋ/ may underlyingly be /ng/.

5. Closing Remarks

Many of the points we have made in this paper admittedly require further refinement and validation. However, we have shown quite clearly that weight is definitely a major factor in linguistic structuring. Although our discussion here has been confined to English, the weight variable very likely is also of relevance to the description and explanation of any other human language. Thus it is very likely that weight is a universal linguistic variable to be reckoned with in any complete and valid theory of language.

As we see it, research on weight as a linguistic variable can shed a great deal of light on a wide range of linguistic phenomena. It will especially be instrumental in explaining various phenomena relating to stress, syllable structure, phonotactics, reduction, and movement. No explanation of these phenomena would be complete without some recourse or other to the weight variable. This calls for intensive future research on this variable.

As has been noted in connection with each of the four principles discussed in the present study, weight is a variable of crucial relevance to language processing. Thus the study of the weight variable can also be justified by the light it promises to throw on the nature of human communication by linguistic means.

We may also point out here that there appears to be a fairly substantial intersection between weight and distance as linguistic variables. This may follow from the fact that linguistic weight increases in proportion to referential distance. We are suggesting that weight and distance as linguistic variables are each definable at least in part in terms of the other. A detailed discussion of the distance variable is available in Park (1980, 1981b, 1982, 1983 and 1984).

One of the things about the weight variable that we have not discussed in any detail in this paper has to do with the fact that it often leaves ortho-

graphic traces behind. An interesting example is to be found in the German practice of beginning every noun with an uppercase letter in apparent recognition of its heavy information weight vis-à-vis other parts of speech. Another interesting example has to do with English acronyms such as the following.

(303) a. NA (=not applicable)
b. MT (=megaton)
c. TM (=trade mark)
(304) a. MS or ms (manuscript)
b. P/C or p/c (=petty cash)
c. AA or aa (=author's alteration)

As can be seen from (303) and (304) above, English acronyms tend to be capitalized even when they do not stand for proper nouns. This is because they stand for longer source expressions so that they are quite compact and heavy. Phenomena of this kind, which are in line with the "weight-over-size" principle, are treated in some detail in Park (in preparation).

Note in this connection that the weight-over-size principle apparently throws interesting light on the peculiar practice of capitalizing the first-person singular pronoun in English. *I* is short for its Middle English equivalent *ich* so that it carries the weight of not just the letter *i* but the other two letters, i.e. *c* and *h*. Thus the pronoun *I* may be written with an uppercase letter under the pressure of the weight-over-size principle.

Before we conclude, four caveats may be in order. Firstly, we have ignored in our discussion minor changes in meaning that may accompany changes in form on the assumption that they are not of crucial relevance. Secondly, the application of our four principles is not an either-or proposition but rather a matter of relative degree. Thus a given phenomenon is in or out of line with a given principle only in relative terms, never in absolute terms. Thirdly, other variables may override or nullify the weight variable, as we have suggested in connection with (242), (243), and (253) in 3.7. Fourthly, we have not done sufficient justice to the fact that weight offen bears on an expression in more than one way.

Regarding the fourth caveat, let us consider the coordinate noun phrases in (305) below.

(305) a. *Men and women* have been farding since time immemorial.
b. *Women and men* have been farding since time immemoiral.

Both coordinate noun phrases here are perfectly natural. However, *women and men* may be (slightly) more natural here than *men and women.* For *women* is less marked and thus less heavy than *men* in the context of farding. That is, as subject of *farding, women and men* is more in line with the weight-over-order principle than *men and women* so that (303b) is more natural than (305a).

Ordinarily, however, *men* is less marked and thus less heavy than *women* so that *men and women* is more in line with the weight-over-order principle than *women and men.* Besides, *mén ănd wómĕn* is more in line with the weight-over-rhythm principle than *wómĕn ănd mén.* Thus *men and women* is ordinarily more natural than *women and men.*

From examples such as (305), we can learn that weight often applies to linguistic expressions non-uniquely so that it allows two or more alternatives that appear to be mutually contradictory. By way of bringing our discussion here to a close, we may point out that the coexistence in English of *pepper-and-salt* and *salt-and-pepper* is just another example ascribable to the non-unique application of the weight variable.

Extensive as our discussion of the weight variable has been, we may merely have scratched the surface. We simply have not had the space to include numerous phenomena of relevance such as that illustrated by the following data.

(306) a. one thousand (=1,000)
b. *thousand one (=1,000)
(307) a. seven million (=7,000,000)
b. *million seven (=7,000,000)

Note here that *one* and *seven* are numerically smaller and thus lighter than *thousand* and *million* respectively so that (306a) and (307a) comply with the weight-over-order principle while (306b) and (307b) do not. Only in a separate paper of considerable length can we hope to do justice to such weight-governed phenomena as have been left out of consideration here.

REFERENCES

Jespersen, Otto (1909–1949) *A Modern English Grammar on Historical Principles I–VII,* Munksgaard, Copenhagen.

Marchand, Hans (1969) *The Categories and Types of Present-Day English Word-Formation,* 2nd ed., Beck, Munich.

Park, Nahm-Sheik (1977a) *Variables in Sentential Readability–with Special Reference to EFL/ESL for Korean Learners,* Georgetown University Ph. D. Dissertation.

_____ (1977b) 'Structural Density as a Comprehensibility Variable,' *Language Research* 13.2, 95–103, Seoul National University, Seoul. (In Korean)

_____ (1980) 'The Distance Variable in Linguistic Description with Special Reference to English,' *Languages and Linguistics* 6, 9–32, Hankuk University of Foreign Studies, Seoul. (In Korean)

_____ (1981a) 'Weight as a Linguistic Variable,' an unpublished paper read at the annual spring seminar of the Linguistic Society of Korea.

_____ (1981b) 'Distance and Length in Language with Special Reference to English,' an unpublished paper read at a seminar cosponsored by the East-West Center and the University of Hawaii Department of Linguistics.

_____ (1982) 'Constraint on Proximate Repetition with Special Reference to English,' *Language Research* 18.1, 217–231, Seoul National University, Seoul.

_____ (1983) 'Proximity and Reduction,' *Language Research* 19.1, 75–102, Seoul National University, Seoul.

_____ (1984) 'The Constraint on Proximate Repetition and Morphophonological Explanation in English,' *Language Research* 20.4, 367–415, Seoul National University, Seoul.

_____ (1985) 'Natural Word Order in English,' *English Studies* 9, 11–32, Seoul National University, Seoul. (In Korean)

_____ (In Preparation) 'The Law of Conservation and Linguistic Explanation.'

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, Jan Svartvik (1985) *A Comprehensive Grammar of the English Language,* Longman Group, London.

Department of English
Seoul National University
56–1 Shinlim–dong Kwanak–gu
Seoul 151–742
KOREA