

The Law of Conservation and Linguistic Explanation*

Nahm-Sheik Park

The main thesis of this paper is that the law of conservation operates in human language in that linguistic elements tend to have their inherent quality conserved over a fairly long stretch of time. For example, deleted linguistic elements tend to leave behind their traces, which are generally proportionate to their original quality. Although our thesis is validated with examples from English only, we assume that the law of conservation is a universal linguistic tendency.

0. Introduction

In this paper, we will show that something like the law of conservation (of mass) is often found in human language, which sheds light on many interesting and otherwise baffling linguistic phenomena. Although all of our evidence is from English, we assume that the law of conservation is a universal linguistic tendency and is thus applicable as such *mutatis mutandis* to other languages as well.

For convenience of exposition, this paper is divided into three main parts. The first part, called Compensatory Weighting, deals with the fact that linguistic reduction is usually compensated for in one way or another. The second part, called Weight Equilibrium, shows that a linguistic element tends to have its canonical weight (or mass) assigned in such a way that the limits of this canonical weight (or mass) may not be violated. The third part, called Surface Manifestation of Underlying Quality, makes the point

* The present study was supported in part by a 1991–1992 grant from the Seoul National University Development Fund. I am grateful to Hee-Rahk Chae, Choon-Hak Cho, James Foley, Ki-Sun Hong, Hyok-Seung Kwon, Young-Suk Lee and Dae-Sik Min for their comments on the pre-publication draft of this paper. Any errors or shortcomings that remain are entirely my responsibility.

that the underlying quality of a linguistic element tends to be manifested on the surface in ways that are commensurate with that underlying quality.

1. Compensatory Weighting

As a general rule, a reduced linguistic element leaves a trace behind in compensation for the reduction (Park 1981a and 1989). The trace here can take a number of different forms, which may be grammatical, phonological or orthographic in nature.

1.1. Grammatical Compensation

The definite article often serves as a trace compensating for a portion deleted from a noun phrase/clause, as can be seen from the following examples.

- (1) a. *the* aged (\leq aged *people*)
- b. *the* refined (\leq *what is* refined)
- c. *the* Hill (\leq *Capitol* Hill)
- d. *the* Yard (\leq *Scotland* Yard)

The definite article here compensates for the loss of *people*, *what is*, *Capitol* and *Scotland* in (1a), (1b), (1c) and (1d) respectively. Incidentally, the arrow here and in the rest of this paper is used in such a way that $X \Rightarrow Y$ means “Y derives from X” while $X \leq Y$ means “X derives from Y.”

Note that we may similarly account for the common practice of using the definite article with nationality adjectives in sibilants, e.g. *the Japanese* in the sense of ‘Japanese people in general’. The definite article here may be interpreted as a trace for ‘people in general’ or something of the sort.

The definite article in the following two sentences also serves as a trace for the deleted noun heads.

- (2) a. Relations between North *Korea* and *the* South are improving.
- b. New *dogmas* are replacing *the* old.

The definite article in (2a) and (2b) above is a trace compensating for *Korea* and *dogmas* respectively. Note in this connection that *National Geo-*

graphic, the name of a monthly magazine, is sometimes shortened to *The Geographic*. Needless to say, the definite article here apparently serves to compensate for the loss of *National*.

It is interesting to note in this connection that *influenza* may be reduced to *flu*, which is often preceded by the definite article. This optional definite article here arguably serves to compensate for the part deleted from *influenza*, i.e. for *in-enza*.

Just as the definite article often compensates for a portion deleted from a noun (phrase), the particle *up* often compensates for a portion deleted from a verb (phrase). We may cite the following examples in support of this contention.

- (3) a. Open *up*. (<=Open *the door*.)
 b. Shut *up*. (<=Shut *your mouth*.)
 c. Buckle *up*. (<=Buckle *your seatbelt*.)
 d. Light *up* (<=Light *your cigarette*)
 e. Did he show *up*? (<=?Did he show *himself*?)

The particle *up* in each sentence of (3) is a trace for the deleted object noun phrase. It could also compensate for other types of verb-phrase elements, as in *Speak up*, in which *up* could arguably be a trace for *louder* or *your mind* depending on the meaning intended. The expressions *team up*, *line up* and *pair up* arguably derive from *form a team*, *form a line* and *form a pair* respectively so that *up* here may be said to compensate for the deleted verb *form*, whose main function here is to convert *team*, *line* and *pair* from noun to verb. Note here that *up* in *suit up* and *dress up* also serves to convert *suit* and *dress* from noun to verb. Parenthetically, other particles may sometimes function as traces for (mostly adverbial) portions deleted from verb phrases.

The particle *up* in the following sentence is arguably a trace for the part deleted from *confess*.

- (4) Why don't you *fess up*? (<=Why don't you *confess*?)

Our discussion here throws light on the fact that *divvy* is usually followed by *up*. In our terms, *up* in *divvy up* arguably compensates for the portion deleted in the derivation of *divvy* from *divide*.

Suffixes also serve as traces for deleted linguistic elements. For example,

-*ly* often compensates for portions deleted from superordinate copular clauses, as can be seen from the following examples.

- (5) a. Evidently, he is to blame. (<=*It is evident that* he is to blame.)
 b. More importantly, she is eminently qualified. (<=*What is more important, she is eminently qualified.*)

The suffix -*ly* here may be said to compensate for the deleted portion *It is ... that* or for *What is ...*. As we will see in 1.2 below, the commas in the two sentences of (5) above also constitute part of the trace for the deleted portions.

The quasi-plural suffix -*s*, too, is sometimes used as a compensatory trace for reduction from nouns or noun phrases. Let us consider the following examples.

- (6) a. turps (<=*turpentine*)
 b. tails (<=*tailcoat*)
 c. regimentals (<=*regimental uniform*)
 d. academicals (<=*academic(al) costume*)
 e. leftovers (<=*leftover food*)

The suffix -*s* here is arguably a trace for -*entine*, -*coat*, *uniform*, *costume* and *food*, i. e. the portions deleted in the derivation of *turps*, *tails*, *regimentals*, *academicals* and *leftovers* from *turpentine*, *tailcoat*, *regimental uniform*, *academic(al) costume*, and *leftover food* respectively.

Admittedly, we may also argue that the -*s* in *tails*, *regimentals* and *academicals* is not only a compensatory trace but also a bona fide plural suffix. For *tails*, for one, implies two tails so that it is semantically plural, so to speak. The reduction of *mathematics* to *maths* in British English may involve a similar suffix. We will have more to say concerning names of sciences ending in -*ics*, as in *mathematics* and *linguistics*, in connection with (194) in 3.3.

Similar in nature is the quasi-plural suffix -*s* in such words as *finals* (for *final game or examination*) and *comprehensives* (for *comprehensive examination*). In our terms, the quasi-plural suffix here serves to compensate for the loss of *game*, *examination* or the like. Similarly explainable is the suffixes in *undies* (for *underwear*).

The quasi-plural suffix -*s* sometimes combines with the definite article to

compensate for deletion, as can be seen from the following paraphrase pair.

- (7) a. She is the queen of *figure skating* in Korea today.
 b. She is the queen of *the figs* in Korea today.

The *-ing* suffix often compensates for a conjunctive structure especially for one with a subordinate conjunction. The following sentence contains a case in point.

- (8) Weather *permitting*, we are going on a picnic this afternoon. (<=*If the weather permits*, we are going on a picnic this afternoon.)

Note at this point that an accusative nominal is arguably heavier than its nominative counterpart in that it is a condensation of a nominative nominal plus something else, i.e. the objective case. Along this line of reasoning, we may say that *me* is a condensation of *I am* in the following sentence.

- (9) He's taller than *me*. (<=*He's taller than I am*.)

Note here that *He's taller than I* is less natural than (9) for the reason that *I* alone is not heavy enough to stand for *I am*.

Likewise, *me* in the following exchange may be said to be heavier than *I* since it is a condensation of *I* and something else.

- (10) A: Who's there?
 B: It's *me*, John. (= *Me*, John.)

It's me or *Me* here may be said to be short for the underlying *It's I that am here* so that it contains more material than just (*It's*) *I*.

1.2. Punctuational Compensation

Punctuation marks often serve as a trace for deleted linguistic elements, as is exemplified by the following data.

- (11) a. Prof. Young (<=*Professor Young*)
 b. Ag. Econ. (<=*Agricultural Economics*)
 c. a. k. a. (<=*also known as*)

The periods after the clipped forms here help to compensate for the portions deleted from the words shown in the parentheses.

A comma may stand for a conjunction or a preposition, as is shown in the examples below.

- (12) a. X₂, Y and Z (<=X *and* Y and Z)
 b. X₂, Y or Z (<=X *or* Y or Z)
 c. The problem is₂ he is not strong enough. (<=The problem is *that* he is not strong enough.)
 d. I'm so good₂ I can do it with my eyes closed. (<=I'm so good *that* I can do it with my eyes closed.)
 e. He lives in Seoul₂ Korea. (<=(?)He lives in Seoul *in* Korea.)

A comma often compensates for a longer stretch of material than just a conjunction. The commas in (13) below are cases in point.

- (13) a. *True*₂ he is a genius. (<=*It is true that* he is a genius.)
 b. *Evidently*₂ he is a genius. (<=*It is evident that* he is a genius.)

It is interesting that the comma is optional in (14a) below while it is obligatory in (14b), which is a transform of (14a).

- (14) a. Though *he was suffering* from arthritis₂ he entered the marathon.
 b. Though *suffering* from arthritis, he entered the marathon.

We may argue here that the comma is obligatory in (14b) because it has to compensate for the deleted portion, i.e. *he was*, and is thus fairly heavy.

An apostrophe often serves as a trace for the part deleted from a linguistic element, as can be seen from the examples below.

- (15) a. the '90s (<=the 1990s)
 b. ne'er (<=never)
 c. He isn't here. (<=He is *not* here.)

The colon or semicolon serves as a compensatory trace for a comma plus some conjunctive element that follows, as can be seen from the following examples.

- (16) a. I have some good news for you₂; you will get the job. (<=I have some good news for you₂ *which is that* you will get the job.)
 b. Those who lead must be considerate₂; those who follow must be responsive. (<=Those who lead must be considerate, and those

who follow must be responsive.)

We can say here that the (semi)colon is informationally and thus formally heavier than the comma, which accords with the law of conservation.

1.3. Phonological Compensation

Originally weak phonological elements tend to get fortified when they carry not just their own weight but also the weight of such elements as have been deleted. Let us take the following clipped words for example.

- (17) a. *rehab* (<=*rehabilitation*)
 b. *recap* (<=*recapitulation*)

In both (17a) and (17b) here, the nucleus of the second syllable gets elevated from the unstressed schwa in the parenthesized word to the more prominently stressed /æ/ in its clipped derivative. In our terms, this fortified /æ/ in the clipped word serves to compensate for the syllables deleted from the longer word. Needless to say, similarly explainable is the upgrading of the second-syllable nucleus in such clipped words as *info* and *intro* derived from *information* and *introduction* respectively. Note here that the nucleus of not only the second syllable but also the first syllable gets upgraded to /ow/ in *promo* derived from *promotion*.

We can observe a similar compensatory trace in the second syllable of either clipped word below, in which rhyme is apparently also a factor.

- (18) a. *sci fi* (<=*science fiction*)
 b. *hi fi* (<=*high fidelity*)

Compensatory weighting is instrumental in accounting for the derivation of the following words from their parenthesized sources.

- (19) a. *parlous* (<=*perilous*)
 b. *frantic* (<=*frenetic*)
 c. *Bedlam* (<=*Bethlehem*)

Two weaker vowels combine to give rise to a stronger vowel in the historical shortening process exemplified by each pair in (19) above, where the stronger vowel may compensate for the reduction involved here. The plosive /d/ in (19c) is an upgraded version of the fricative /θ/ and as such

helps to compensate for the reduction involved in the derivation of *Bedlam* from *Bethlehem*. Note in this connection that the monosyllabic prefix *mi-* in *migraine* derives from the disyllabic *hemi-* and is pronounced with the strong nucleus /ay/ so as to compensate for its disyllabic origin.

Compensatory weighting also throws light on the diphthongal nucleus in the first syllable of *halfpenny* as well as in the common British pronunciation of the male given name *Ralph*. We may note in this connection that the substandard *don't* for *doesn't* owes its strong diphthongal nucleus to compensatory weighting.

As can be seen from the examples given in the above paragraph, the nucleus of a syllable often takes the form of a diphthong (rather than a monophthong) to compensate for the loss of a consonant in the coda of the same syllable. Consider additional examples such as *yeoman* (cf. *young man*), *tithe* (cf. *tenth*) and *tooth* (cf. *-dont-* in *orthodontics*).

Consonants also get similarly fortified to compensate for deleted segments, as in the following examples.

- (20) a. *bike* (<=*bicycle*)
 b. *Frank* (<=*Francis*)

The lighter *c*, i.e. /s/, is replaced here by the heavier *k*, i.e. /k/, when *bicycle* and *Francis* are shortened to *bike* and *Frank* respectively. Needless to say, the heavier *k* here stands for the deleted portions *-cycle* and *-cis* respectively, both of which begin with the lighter *c*. Parenthetically, the weight (=fortisness) of a consonant is in inverse proportion to its sonority so that stops are heavier than non-stops, for example.

The kind of compensatory weighting under discussion here is especially productive in the shortening of first names such as that observed in (20b). Let us consider the following examples.

- (21) a. *Hal* (<=*Harry*)
 b. *Lal* (<=*Larry*)

Note that /l/ is a (less sonorous and thus) heavier liquid than /r/ so that *-l* here may be said to compensate for *-rry*. We may cite the following additional examples.

- (22) a. *Ted/Ned* (<=*Edward*)
 b. *Dick/Rick* (<=*Richard*)

- c. *Bob* (<=*Robert*)
- d. *Bill* (<=*William*)
- e. *Betty/Betsie* (<=*Elizabeth*)

T- and *N-* compensate for *-ward* in (22a), (*D*)-*ck* for *R-chard* in (22b), *B-* for *R-ert* in (22c), and *B-* for *W-iam* in (22d). We may note in connection with (22e) that the hypocoristic suffix *-y* or *-ie* combines with /t/ or /ts/ to compensate for the shortening involved.

Admittedly, *Edward*, *Richard*, *Robert* and *William* may also be abbreviated to *Ed*, *Rich*, *Rob* and *Will* respectively with no apparent compensatory trace for the portions deleted. Other things being equal, however, /ed/, /ritʃ/, /rab/ and /wil/ are phonetically longer in *Ed*, *Rich*, *Rob* and *Will* than in *Edward*, *Richard*, *Robert* and *William* respectively. Thus this greater phonetic length arguably compensates for the abbreviation involved.

Incidentally, there is reason to believe that in examples like (17)–(22) the fortified sound used as a compensatory trace is usually not just any sound. On the contrary, it is a sound closely related in some way to part of the deleted portion for which it stands. In (22d), for one, /b/ in *Bill* is clearly related to /w/ in *William* in that both /b/ and /w/ are bilabial sounds with the former being a fortified version of the latter, so to speak.

Compensatory weighting helps explain why auxiliary or quasi-auxiliary verbs may not be contracted clause-finally, as we can see from the paraphrase pair below.

- (23) a. He's taller than *I am*.
- b. *He's taller than *I'm*.

We may argue here that the clause-final *am* in (23a) is a condensation of the underlying *am tall* and thus has to be heavy enough to compensate for the missing *tall* so that it is too heavy to allow contraction to *'m*.

Compensatory weighting also helps explain why one and the same word is assigned heavier stress as an adverbial particle than as a preposition, as is the case with the two tokens of *in* in the following pair of sentences.

- (24) a. He was *in* the house.
- b. He was *in*.

Given proper context the two sentences here could be paraphrases of each other so that *in* in (24b) may be considered a condensed form of *in the*

house in (24a). In fact, the particle *in* is always understood as a condensation of the preposition *in* plus the trace of its complement. Thus *in* is normally pronounced with heavier stress in (24b) than in (24a) because it is underlyingly heavier in (24b) than in (24a).

It is interesting that a phrase-final preposition is pronounced more prominently than its phrase-initial counterpart so that *to* gets heavier stress in (25b) than in (25a) below.

- (25) a. They spoke *to* John.
 b. John was spoken *to*.

Although *John* is the prepositional complement in both sentences here, the phrase-final *to* in (25b) is heavier than normal because it is short for itself plus the trace of *John* while this is not the case with *to* in (25a). This is why *to* is pronounced with heavier stress in (25b) than in (25a).

Incidentally, Shakespearean English affords us an interesting example of relevance to our discussion. Let us consider the following sentence from *The Merchant of Venice*.

- (26) Which *of* my ships art thou the master *off*?

In Shakespearean English, *of* and *off* were two variants of the same morpheme with the latter being the heavier variant. This difference in weight is indicated orthographically by the use of one more *f* in *off* than in *of*. It is also indicated phonologically by a stronger pronunciation for *off* than for *of*. Note that the heavier *off* here carries its own weight plus that of the trace of its complement, i.e. *which*, while this is not the case with the lighter *of*.

Given our discussion here, it should come as no surprise that the phrase-final *to* in the italicized phrases below are rather prominently stressed.

- (27) a. We all *set to* and got the job done in no time.
 b. The committee *turned to* and produced a plan.
 c. When I *came to*, I was in a hospital.
 d. How did you *bring him to*?

In all the phrases in question here, the preposition *to* carries its own weight plus that of the trace of its missing complement so that it is heavier here than is ordinarily the case.

1.4. Orthographic Compensation

Various types of linguistic shortening are compensated for by orthographic means also. The frequent use of uppercase letters for acronyms is an interesting mode of compensation. Let us consider the following examples.

- (28) a. KO'd (<=knocked out)
 b. DJ's (<=disc jockeys)
 c. VOA (<=Voice of America)
 d. R & D (<=research and development)
 e. MIA (<=missing in action)
 f. POW's (<=prisoners of war)
 g. STD's (<=sexually transmitted diseases)
 h. I (<=(Middle English) ich)

The capital letters in the abbreviations here are supposed to compensate for the portions deleted from the original words in the parentheses. We may note that VOA has a variant in *VoA*, in which *o* is written with a lowercase letter to reflect the fact that it stands for the relatively light, insignificant function word *of*. An exactly identical explanation applies to the abbreviation of *Museum of Modern Art* to either *MOMA* or *MoMA*. We may also note that the inflectional suffixes on the abbreviations above are rendered in lowercase letters because of their lack of inherent weight.

Quasi-acronyms such as the following also involve the use of capital letters as a means of compensating for the deletions involved.

- (29) a. TV (<=television)
 b. TB (<=tuberculosis)
- (30) a. CA (<=California)
 b. TX (<=Texas)
- (31) a. A-bomb (<=atomic bomb)
 b. H-bomb (<=hydrogen bomb)

We may point out here that *atomic bomb* or *hydrogen bomb* may be abbreviated to *Bomb* with a capital *B*, which also vindicates our claim that abbreviations are often compensated for by means of capital letters.

Block language exemplifies another case of recourse to capital letters to

compensate for linguistic reduction. Book titles, chapter headings, table of contents listings, etc. normally have their initials capitalized so as to show that they stand for whole blocks of information. This is of course another piece of evidence for our claim that linguistic elements almost always leave their traces behind.

Our last example of orthographic compensation for a deleted linguistic element is provided by the derivation of *till* from *until*. It is worthy of note here that *till* ends in two tokens of *l* while *until* ends in just one token of the same. The additional token of *l* in *till* here is arguably the compensatory trace for the deleted *un-* of *until*.

2. Weight Equilibrium

A linguistic element of a given (functional) category is assigned more or less constant weight specific to that category such that its derivational changes have to satisfy the constraints of that category-specific canonical weight (Park 1981a and 1989). All the examples of compensatory weighting discussed in the preceding chapter may be cited in support of this contention. In fact, what we may call a principle of weight equilibrium prevails, whereby a heavy constituent alternates with a light one and vice versa to guarantee weight equilibrium for members of a given category.

2.1. Phonological Weight

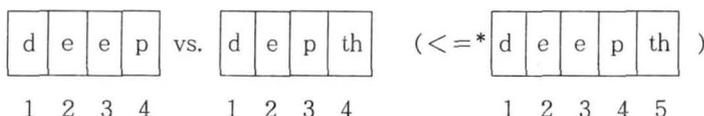
The phonological weight assigned to a syllable, especially a stem-core syllable, is such that a heavy nucleus normally may not be followed by a heavy consonantal coda. This is evidently designed to prevent the syllable from *getting overweight* by *exceeding the upper limits of its category-specific canonical weight*. Let us compare the two members of each of the following word pairs with specific reference to their nuclei and codas.

- (32) a. deep vs. depth (<=*depth)
 b. keep vs. kept (<=*keep<=*kept)
 c. five vs. fifth (<=*fayfθ/)

The parenthesized forms here, in which a heavy nucleus is followed by a

heavy consonantal coda, exceed the limits of the weight allowed a normal syllable in English. To resolve this problem, the originally heavy nucleus in each pair here is downgraded to a lax vowel in front of the heavy consonantal coda in the second member of the pair. In this way, we can maintain a balance of phonological weight between the first and second words in each pair.

Incidentally, the balance of weight achieved by (32a), for one, may be graphically represented as follows.



We can see here that both *deep* and *depth* are assigned four spaces (or units) of weight so that there is clearly an equilibrium of weight between the two words achieved through the variation of the nucleus and the coda. Needless to say, the two tokens of *e* in *deep* above represent the “onset” /i/ and the “coda” /y/ for the diphthongal nucleus /iy/ of this word. Note that the underlying form **depth* for *depth* is overweight by one “segment”.

The vowel alternation exemplified by each of the following word pairs may be accounted for along similar lines.

- (33) a. wide vs. width
 b. broad vs. breadth
 c. long vs. length

The vowel nucleus in the first member of each pair here is diphthongal while its counterpart in the second member is monophthongal. On the other hand, the consonantal coda in the first member of each pair is monoconsonantal while its counterpart in the second member is biconsonantal. Thus it is the case that a heavy nucleus is followed by a light coda in the first member of each pair here while a light nucleus is followed by a heavy coda in the second member. Again this is apparently designed to maintain a balance of weight between the two members of each pair here.

We can provide a similar account for the vowel alternation exemplified by each of the following pairs.

- (34) a. mind vs. mint
 b. hind vs. hint
 c. wild vs. wilt
 d. mild vs. milt

Since a voiceless consonant is (less sonorous and thus) heavier than its voiced counterpart, /t/ is heavier than /d/. Thus we can say that in each pair here the heavier nucleus /ay/ combines with the lighter coda /nd/ while the lighter nucleus /i/ combines with the heavier coda /nt/. This is again in keeping with the principle of weight equilibrium.

The stem-vowel alternation exemplified by each of the following word pairs is also motivated by a desire for weight equilibrium.

- (35) a. sane vs. sanity
 b. serene vs. serenity
 c. crime vs. criminal
 d. bone vs. bonfire
 e. know vs. knowledge
 f. house vs. husband

Note that the concept of the coda for the stem syllable here may be extended to include whatever follows the nucleus. Under this conceptual extension, the coda for the stem nucleus in the second member of (35a), for one, would be *-nity*, which would count as a heavy coda on account of its (poly) syllabicity. Given this interpretation, the stem-vowel alternation in (35) can be explained in exactly the same way as that in (32)–(34).

The productivity of the “suffixal” element that forms part of the coda in examples like (35) is apparently of relevance to the operation of weight equilibrium. As a general rule, the productivity of the “suffixal” element in question here is in inverse ratio to the weight it contributes to the coda. Thus the noun suffix *-ity*, for one, is less productive and thus heavier than the noun suffix *-ness*, which seems to be the main reason why the stem vowel remains tense in *saneness* while it becomes lax in *sanity*. Similarly explainable is the difference in first-syllable nucleus between *post-war* and *posthumous*, between *knowing(ly)* and *knowledge*, between *heals* and *health*, etc. It may very well be the case that a productive “suffixal” element is not as closely bound to the stem as its less productive counterpart so that the former lends less weight to the coda in question than does the latter.

A more productive suffix is apparently preceded by a heavier boundary than is its less productive counterpart. Thus, for example, *-ness* apparently follows a heavier boundary than does *-ity*. If this is indeed the case, then we may argue that the heavier boundary preceding *-ness* bars the suffix from lending its full weight to the coda of the stem-core syllable whereas the lighter boundary preceding *-ity* does not do so. This boundary argument casts plenty of light on why *saneness* has a tense vowel while *sanity* has a lax one.

It also appears to be the case that productive elements are generally younger than their less productive counterparts. Given this phenomenon and our discussion in the immediately preceding paragraph, we may speculate here that younger elements tend to be more immune to the pressure of weight equilibrium than older ones are.

There is plenty of evidence that it apparently takes a considerable period of time for weight equilibrium to fully assert itself. In word pairs involving such relatively late formations as *comatose* (from *coma*), the vowel alternation may not be as well established as in (35). However, the first vowel here is more frequently monophthongal than it is diphthongal in *comatose* while it is always diphthongal in *coma*, which apparently indicates that even here the force of weight equilibrium is very much in evidence.

Although *chastisement* is evidently not quite as recent a formation, something similar can be said about the pair *chastise/chastisement*, in which the letter *i* is either /i/ or /ay/ in the second member while it is always /ay/ in the first. The third-syllable vowel alternation between *advertise* and *advertisement*, especially in British English, also bears the unmistakable hallmark of the principle of weight equilibrium.

Weight equilibrium is evidently at work in cognate triads such as the following.

- (36) a. divide vs. divisive vs. division
 b. decide vs. decisive vs. decision

The coda-*sive* is arguably lighter than the coda *-sion* in that the former is monosyllabic while the latter is (diachronically and thus) underlyingly disyllabic, i.e. *-si-* + *-on*. We may thus contend that weight equilibrium is maintained here with a diphthongal vowel before the lighter coda, as opposed to a monophthongal vowel before the heavier coda. As we shall see in

connection with (48), the voicing alternation of the sibilant segment between the second and third members of each triad here is also a consequence of the pressure of weight equilibrium.

Note in this connection that *divisive* is sometimes pronounced /divisiv/ with the monophthongal vowel /i/ stem-finally. This may be due to the fact that the coda is slightly heavier in *divisive* than in *divide* so that it may be felt necessary to maintain weight equilibrium between the two words by lightening the stem-final vowel in front of the relatively heavier *-sive*.

Let us now turn our attention to the interesting fact that the suffix *-ion* induces the laxing of the stem-final nucleus in the first member of either pair below while it does not do so in the second member of the same pair.

- (37) a. provision (<=provide + -ion)
 b. invasion (<=invade + -ion)
 (38) a. ignition (<=ignite + -ion)
 b. vacation (<=vacate + -ion)

Note that /ay/ is heavier than /ey/ so that the (already heavy) suffix *-ion* may exert heavier pressure on the former to reduce than on the latter. This is arguably why /ay/ gets laxed to /i/ in the first member of either pair above while /ey/ does not get similarly laxed in the second member of the same pair.

The quest for weight equilibrium also underlies the vowel alternation exemplified by the following cognate pairs.

- (39) a. car vs. carry
 b. bar vs. barrier

We can see here that the weight of the (stem-final) vowel is in inverse proportion to that of the coda, which is evidently a consequence of the search for weight equilibrium between the two members of each pair here.

Incidentally, the stem-final vowel alternation between *portray* and *portraiture*, i.e. between /ey/ and the schwa, may be similar in nature to that exemplified by (39) above. Note also that the stem-final syllable in *portrait* has a coda halfway in weight between *portray* and *portraiture*, which may explain why its nucleus alternates between the light schwa and the heavy /ey/. Note further that the stem-final vowel in *portrayal* is diphthongal (and thus heavy) despite the (mono)syllabic (and thus potentially heavy)

coda. This may be because the coda here is lighter than the (mono)syllabic (and thus heavy) coda in *portraiture* and lighter than even the subsyllabic coda in *portrait*. There may also be other contributory factors involved here. For example, *-al* may be preceded by a major intermorphemic boundary which keeps it from pressuring the stem-final vowel to get reduced to the schwa or some such weak vowel.

Note the first-vowel alternation, especially in British English, for the following pairs of cognates.

- (40) a. drama vs. dramatist (/ɑ:/ vs. /æ/)
 b. class vs. classic (/ɑ:/ vs. /æ/)

The alternation in question here is due to the difference in weight between the two codas, i.e. between *-ma* and *-matist* in (40a) and between *-ss* and *-ssic* in (40b). Other heavy codas than *-matist* and *-ssic* here, e.g. *-tize*, *-ssical* and *-ssicist*, apparently exert similar, albeit weaker, pressure on the original /ɑ:/ to commonly reduce to /æ/.

Admittedly, there are varieties of English in which the first-syllable nucleus may be either /ɑ:/ or /æ/ for either member of each pair in (40) above. Even in these varieties of English, however, /ɑ:/ is generally more likely for the first member of each pair than for the second while /æ/ is more likely for the second member of each pair than for the first. It goes without saying that this is again arguably due to the pressure of weight equilibrium.

It is interesting that the first-syllable nucleus in both *trauma* and *traumatic* may be pronounced with either the clearly diphthongal /aw/ or the less diphthongal /ɔ:/ with the proviso that /aw/ is more frequent with *trauma* and /ɔ:/ with *traumatic*. The alternation here is of course ascribable to the difference in the weight of the coda between *-ma* and *-matic*. Here again other heavy codas than *-tic*, e.g. *-tize* and *-tism*, apparently exert similar, albeit weaker, pressure on the original strong diphthong to reduce to a lesser diphthong.

A similar account is applicable to the interesting fact that the stem vowel is less likely to be diphthongal in *cyclical* than in *cyclic* while it is always diphthongal in *cycle* and always monophthongal in *encyclical*. Note that the weight of the vowel nucleus in question here is in inverse ratio to that of the rest of the word, which is in line with the principle of weight equilibri-

um.

(41) below offers an interesting historical example of coda alternation resulting from the pressure of weight equilibrium.

- (41) a. rough/cough/laugh
 b. dough/bough/plough

The word-final *-gh* here, which used to be pronounced /x/, has diverged into /f/ in words of the (41a) type and into zero in words of the (41b) type. Note that the nucleus is monophthongal when the coda is non-zero, i.e. /f/, while it is diphthongal when the coda is zero. Thus a simpler nucleus is followed by a less simple coda and vice versa, which is basically in line with the principle of weight equilibrium.

Let us consider the well-known voicing alternation exemplified by the following singular-plural pairs.

- (42) a. wife vs. wives
 b. mouth vs. mouths

The nucleus is diphthongal here so that the addition of the plural suffix to the singular form would result in a situation where a heavy nucleus would be followed by a heavy coda, i.e. by a cluster of voiceless consonants. As a means of helping resolve this situation, we resort to voicing the stem-final fricative here and consequently the plural suffix, which results in a considerably lighter coda, i.e. a cluster of voiced consonants. For the weight of a consonantal segment decreases in proportion to its sonority, which in turn rises in proportion to its voicedness.

Note in this connection that the alternation observable in pairs such as *half/halves* is amenable to a similar explanation. It is important to remember that the stem vowel here is long and diphthongal, underlyingly and especially in British English, so that leaving the stem-final consonant voiceless in the plural form would result in a heavy nucleus followed by a relatively heavy cluster. Thus the voicing of the stem-final consonant and consequently of the plural suffix in *halves* is designed to render the word more compliant with the principle of weight equilibrium.

Weight equilibrium is also at work in the suffixal vowel alternation exemplified by the following word pairs.

- (43) a. communism vs. communist
 b. Marxism vs. Marxist
 c. antagonism vs. antagonist

The suffix ends in the lighter coda /zm/ in the first member of each pair here while it ends in the heavier coda /st/ in the second member. The suffixal nucleus here is generally longer (either allophonically or phonemically) before the lighter /zm/ than before the heavier /st/. In fact, the suffixal vowel is often rendered as /i/ before /zm/ and the barred 'i' before /st/.

(44) below offers an interesting historical example of consonant alternation attributable to the pressure of weight equilibrium.

- (44) a. sound/pound
 b. peasant/tyrant

We know that the word-final stops /d/ and /t/, which were not originally part of the words here, were added to sort of round out the word-final codas in question. Note that the voiceless /t/ is heavier than the voiced /d/ and that the lighter /d/ was added to the syllable with the diphthongal (and thus heavier) nucleus /aw/ whereas the heavier /t/ was added to the syllable with the monophthongal (and thus lighter) nucleus /ə/.

The following example may be cited in support of the claim that not just the coda but sometimes the onset may also condition the kind of stem-vowel alternation under consideration here.

- (45) a. Dane vs. Danish
 b. Spain vs. Spanish

The onset is lighter in (45a) than in (45b) in that the former begins with just one (voiced) consonant while the latter begins with two (voiceless) consonants. The onset /sp/ may be heavy enough to combine with the coda *-nish* to lax the original diphthongal nucleus /ey/ to /æ/ in (45b) while the onset /d/ is not apparently heavy enough to do so in (45a). The *read/spread* pair affords us an interesting historical example in support of our argument here. The nucleus, which used to be the same diphthong for both words here, has diverged into the tense /iy/ for *read* and the lax /e/ for *spread*. Note that the lighter onset /r/ precedes the heavier nucleus /iy/ in *read* while the heavier onset /spr/ precedes the lighter nucleus /e/ in

spread, which is in line with the principle of weight equilibrium.

The concept of onset may be extended to include whatever precedes the (syllable) nucleus in question just as the concept of coda may be extended to include whatever follows the nucleus. Recall our discussion of a similar conceptual extension for the coda in connection with (35). The conceptual extension suggested here apparently throws light on the stem-vowel alternation exemplified by the following positive-negative pairs.

- (46) a. pious vs. impious
 b. finite vs. infinite
 c. potent vs. impotent

The diphthongal nucleus in the positive member of each pair here gets laxed to its monophthongal counterpart in the negative member of the same pair apparently under the pressure of weight equilibrium.

Not just vowels but also consonants may sometimes be weakened in order to make room for the weight added by extra material in the coda. Let us take a look at the following examples, where the added weight is accounted for by suffixes.

- (47) a. *divide* vs. *divisive*
 b. *pirate* vs. *piracy*
 c. *electric* vs. *electricity*
 d. *pedagogue* vs. *pedagogic*

Admittedly, the sibilantization exemplified by each pair above is often idiosyncratic in that it is often suffix- or word-specific. It is also due in large part to the fact that the suffix in question begins with an underlyingly high front vowel. The sibilantization here does, however, help reduce the phonological weight of the stem-final consonant when the coda gains weight with the addition of a (syllabic) suffix. This is arguably also due to the pressure of weight equilibrium.

It is interesting to observe in this connection that the voicing alternation exemplified below is also evidently the handiwork of the principle of weight equilibrium.

- (48) a. *divisive* vs. *division*
 b. *evasive* vs. *evasion*
 c. *inclusive* vs. *inclusion*

d. explosive vs. explosion

Note that *-ion* is underlyingly longer and thus heavier than *-ive* in that diachronically the former is dysyllabic while the latter is monosyllabic. Since voicing has the effect of reducing the weight of the segment affected, we may contend here that the voicing alternation exemplified by (48) is due to the pressure of weight equilibrium.

Thus far we have confined our attention almost exclusively to examples of phonemic alternation conditioned by differences in the weight of the coda (and, less often, of the onset as well). However, there are just as many or ever more examples of similarly conditioned allophonic alternation. Let us take the following word pairs for example.

- (49) a. bad vs. bat
 b. lab vs. lap
 c. lag vs. lack

Although the nucleus is phonemically /æ/ in all the words of (49), it is phonetically longer in the first member of each pair here than in the second member. Since a voiceless sound is more fortis than its voiced counterpart, the longer vowel in the first member of each pair precedes a less fortis consonant while the shorter vowel in the second member precedes a more fortis consonant. This accords with the principle of weight equilibrium.

It is interesting that the one and the same vowel phoneme in each word pair below is phonetically longer in the first member than in the second.

- (50) a. hiss vs. hit
 b. his vs. hid
 c. leaf vs. leap
 d. heath vs. heat
 e. breathe vs. breed

The first member of each pair here has a lighter coda than the second since a fricative is lighter than its plosive counterpart. Thus the nucleus here is phonetically longer before a lighter coda than before a heavier one, which is in compliance with the principle of weight equilibrium.

The allophonic alternation of the nucleus exemplified below can also be accounted for in a similar manner.

- (51) a. hall vs. halt
 b. Ben vs. bend
 c. bar vs. bark
 d. miss vs. mist
 e. call vs. called

The first member of each pair here ends in one consonant while the second member ends in that consonant plus another consonant. One and the same nucleus in each pair here is thus phonetically longer in the first member than in the second so that a longer nucleus combines with a lighter coda while a shorter nucleus combines with a heavier coda.

Note at this point the interesting phenomenon of the stem vowel in each of the following triads getting progressively shorter allophonically from left to right.

- (52) a. Ben vs. bend vs. bent
 b. bar vs. bark vs. barked
 c. tide vs. tidy vs. tidily
 d. sick vs. sickly vs. sickness
 e. love vs. lovely vs. loveliness

In each triad here, the first word has a longer nucleus and a lighter coda than the second, which in turn has a longer nucleus and lighter coda than the third. Thus we can see that the phonetic length of the stem vowel here is in inverse ratio to the weight of the coda, which is again explainable as a consequence of the principle of weight equilibrium.

We have thus far confined our attention to vocalic alternation only. There are just as interesting cases of allophonic alternation involving consonants as well. Let us consider the following word pairs with special reference to the phonetic length of the consonantal coda.

- (53) a. fit vs. feet
 b. hop vs. hope
 c. look vs. Luke
 d. tell vs. tale

The nucleus is (phonemically) shorter in the first member of each pair here than in the second whereas the coda is (phonetically) longer in the first member than in the second. Thus one and the same consonantal coda is

phonetically longer after the shorter nucleus while it is phonetically shorter after the longer nucleus. Here again the force of weight equilibrium is clearly in evidence.

Let us take note at this point of such minimal pairs as *take:tack*, *soak:sock* and *like:lick*. The nucleus is heavier in the first member of each pair here than in the second while the coda consonant is (phonetically) slightly heavier in the second member of each pair than in the first. This is all in line with the principle of weight equilibrium. It is noteworthy in this connection that the difference in phonetic weight between the first and second members of each pair here is faithfully reflected orthographically in that the orthographic coda is mono-consonantal in the first member while it is bi-consonantal in the second.

The consonantal alternation illustrated by the following word pairs is also explainable as due to the principle of weight equilibrium.

- (54) a. laid vs. lady
 b. rob vs. robber
 c. rug vs. rugged

Each pair here contains a shared plosive phoneme, which is followed by zero segments in the first member and by one or more segments in the second. Note that this shared plosive is phonetically longer before zero segments, i.e. in the first member, than before non-zero segments, i.e. in the second member, which is again in keeping with the principle of weight equilibrium.

The principle of weight equilibrium sheds light on the aspirated and unaspirated allophones of the plosive phonemes illustrated by the following word pairs.

- (55) a. crew vs. screw
 b. pray vs. spray
 c. tray vs. stray

In each pair here, the first member is one segment shorter than the second in that the first member begins with two consonants while the second begins with the same two consonants preceded by another consonant, i.e. /s/. Thus within the more or less constant limits of lexical weight, the phonetic weight allotted the shared plosive phoneme should be smaller in the second

member of each pair above than in the first. It is precisely for this reason that the shared plosive has its aspiration and voicelessness reduced in the second member. For a segment gains weight in proportion to its aspiration and voicelessness. Parenthetically, it is also for the same reason that /r/ is less devoiced in the second member of each pair here than in the first. All this is of course the work of the principle of weight equilibrium.

Weight equilibrium also affects suprasegmental phonology, as can be seen from the following examples.

- (56) a. black bird vs. blackbird
 b. green house vs. greenhouse

Although both members of each pair here comprise the same two constituent words, the first member ranks higher than the second in that the former is a syntactic phrase while the latter is a compound word. Thus the first member is assigned greater (suprasegmental) weight than the second so that the former has the heavier stress pattern of secondary stress followed by primary stress whereas the latter has the lighter stress pattern of primary stress followed by tertiary stress. There is also a longer juncture between the two immediate constituents in the phrasal combination than in its compound counterpart.

The second immediate constituent of a compound word often has to be drastically reduced (supra)segmentally on account of its meager semantic content (Park 1981a and 1989) and also under the pressure of weight equilibrium. Thus this second constituent often ends up taking the form of a suffix, as is the case with the following words ending in the suffix *-er*.

- (57) a. sleeper (< = *sleeping car*)
 b. diner (< = *dining car*)
 c. washer (< = *washing machine*)
 d. dryer (< = *drying machine*)

In both (56) and (57) above, the second immediate constituent of each compound word (often along with a semantically insignificant part of the first immediate constituent) is reduced suprasegmentally or otherwise so as to enable the compound to fit snugly into the canonical weight confines of the ordinary lexical category.

2.2. Orthographic Weight

The principle of weight equilibrium is also instrumental in explaining a number of otherwise baffling orthographic phenomena. Let us begin our discussion here with the following examples.

- (58) a. *Finn* vs. *Finland* (<=**Finland*)
 b. *Lapp* vs. *Lapland* (<=**Lapland*)
 c. *pass* vs. *past* (<=**passt*)
 d. *all* vs. *already* (<=**already*)
 e. *spill* vs. *spilth* (<=**spilth*)

The first member of each pair here is a monomorphemic word ending in two identical consonant letters, the second of which is dropped when it is followed in the second member by another morpheme beginning with a consonant letter. This dropping may occur in part because the two morphemes, if combined into one bimorphemic word in their original shapes, would somehow exceed the upper limits of canonical orthographic weight assigned to the lexical category. For canonical orthographic weight assigned to a word is largely invariant between monomorphemic and polymorphemic words.

Note incidentally that in each pair of (58) the alternation between the single letter (e.g. *n* in *Finland*) and its geminate cluster (e.g. *nn* in *Finn*) appears to be psychologically real. This is because the single letter here represents a shorter allophone of one and the same phoneme than does its geminate cluster.

Another contributory factor here is the constraint on proximate repetition (Park 1977a, 1977b, 1982 and 1983), according to which a cluster of three consonants or consonant letters is usually simplified by dropping the consonant (letter) in the middle. Parenthetically, *-th* in (58e) is a digraph and thus may be counted as just one letter, i.e. one grapheme. We will discuss in 2.4 another potential variable in the reduction of *all* to *al-* in the derivation of *already* from *all ready*.

The principle of weight equilibrium appears to apply with only very weak force when the second morpheme here is an inflectional suffix realized by a very weak consonant. Let us consider the following example.

(59) *spill* vs. *spilt* (<=**spillt*) vs. *spills* (<=**spils*)

We can see here that the principle of weight equilibrium applies to *spilt*, in which the second morpheme is realized by a hard consonant, while it does not apply to *spills*, in which the second morpheme is realized by a soft consonant. We may say that /t/ here is heavy enough to pressure the stem-final letter *l* to drop while /z/ is not sufficiently heavy to do so. Note in this connection that an inflectional suffix is not bound as closely to the stem as a derivational one, which may be part of the reason why cases like *spills* are immune to the pressure of weight equilibrium while cases like *spilth* are not.

Let us now turn our attention to the following word pairs, which are spelled here the British way.

- (60) a. *honour* vs. *honorary*
 b. *labour* vs. *laborious*
 c. *odour* vs. *odorous*

The spelling *-our*, as in *honour*, normally changes to *-or* when a derivational suffix is added, as in *honorary*, losing the letter *u* in the process. In our terms, this deletion results from the pressure of weight equilibrium and is thus somewhat similar in nature to the loss of one letter in the derivation of the second member of each pair in (58).

Parenthetically, *honourable* retains the letter *u* of *-our* (in British English) despite the relatively heavy suffix *-able*. This may have to do with the fact that *-able* is almost as productive as an inflectional suffix so that it is not so tightly bound to the stem and thus does not exert enough pressure on *-our* to get reduced to *-or*. We are suggesting here that *honourable* retains the letter *u* for exactly the same reason that *honours*, *honoured* and *honouring* do.

Also of relevance to our discussion here is the following example.

(61) *four* vs. *fourteen* vs. *forty*

Fourteen retains *-our* here while *forty* loses *u* so that the former is apparently immune to the principle of weight equilibrium while the latter is not. This is probably because the second morpheme is closer knit to the first in *forty* than in *fourteen* (= *four* # *teen*) so that *-ty* exerts greater pressure on

u to drop than does *-teen*. Incidentally, *fourteen* is historically from *four and ten* while *forty* is close to being a true 'atomic' word, which may be cited in support of our contention here. Incidentally, *fourth* (=4th) is only an apparent exception in that the reduction of *fourth* to *forth* is blocked by the existence of its orthographic twin *forth* (\leftarrow forward).

It is interesting that the first form in each pair below is more frequent than the second in both American and British English.

- (62) a. *glamour* vs. *glamor*
 b. *glamorous* vs. *galmourous*
 c. *galmorize* vs. *glamourize*

We can see here that *-our* is preferred when it is not followed by anything while its reduced form *-or* is preferred when it is followed by a suffix such as *-ous* or *-ize*. This phenomenon, too, is evidently due to the pressure of weight equilibrium.

Similarly explainable is the orthographic reduction of the suffix *-ous* to *-os-*, as exemplified in the derivation of the second member of each pair below from the first.

- (63) a. *generous* vs. *generosity*
 b. *viscous* vs. *viscosity*
 c. *religious* vs. *religiosity*

The reduction involved here is identical in motivation to the reduction of *-our* to *-or-*, discussed in connection with (60) and (61) above, except that it may be slightly more productive. It is interesting that *-ous* does not get reduced to *-os-* before the highly productive suffixes *-ly* and *-ness*, as in *generously/generouslyness*, just as *-our* does not get reduced in British English to *-or-* before the productive suffix *-able*, as in *honourable*. Incidentally, it is apparently always the case that a highly productive suffix follows the major word-internal boundary # so that this boundary keeps the suffix from adding its weight to the stem-final (orthographic) coda.

The following pairs of syllabified words afford us another enlightening example of weight equilibrium operating on the (phono)orthographic plane.

- (64) a. *fin-ish* vs. *fi-nal*
 b. *stud-y* vs. *stu-di-ous*
 c. *leg-ac-y* vs. *le-gal*

d. rad-ic-al vs. ra-di-al

Note that the first syllable ends in a vowel-consonant sequence in the first member of each pair above while it ends in just a vowel in the second member. Note further that the first syllable has a monophthongal nucleus in the first member of each pair here while it has a diphthongal nucleus in the second. Thus we can see that the lighter monophthongal nucleus is followed by a heavier coda, i.e. a consonant letter, while the heavier diphthongal nucleus is followed by a lighter coda, i.e. nothing, which is in line with the principle of weight equilibrium.

2.3. Lexical Weight

There is a general tendency in English to keep words from getting either overweight or underweight, i.e. to keep words within the proper confines of weight assigned to the lexical category. The tendency for words to keep from getting overweight is often manifested by a function word contracted to suffixal form, e.g. *-s* for *is/has/does*, getting more freely attached to a personal pronoun than to a noun. For example, the first sentence in each pair below is slightly more common than the second.

- (65) a. *He's* nice.
 b. *John's* nice.
- (66) a. *She's* done it again.
 b. *Jane's* done it again.

A personal pronoun is a function word and is thus so lightweight that the addition of *-s* would not render it overweight at all. On the other hand, a noun is a content word and is thus inherently weighty so that it may get overweight, albeit ever so slightly, with the addition of *-s*. This is probably why the suffixal *-s* here is not attached to nouns quite as often as it is to pronouns. As we shall see in connection with (75)–(78), contractions like *-s* are too lightweight to qualify as independent words with the result that they end up getting attached to words in their immediate vicinity. This tendency to resist overweighting (as well as underweighting) is a consequence of an effort to comply with the weight limitations imposed on the lexical category.

The tendency to avoid overweight lexical items is also exemplified by the following paraphrase pair.

- (67) a. *I'd have done that.*
 b. (?)*I'd've done that.*

The second sentence here is slightly less natural than the first, at least in written English. This is probably because the double contraction in the second sentence renders a little overweight the word to which it is attached, while the single contraction in the first sentence does not. Another contributory factor here is the constraint on proximate repetition (Park 1977a, 1977b, 1982 and 1983), according to which proximate repetition of similar elements like the two contracted forms in (67b) above is maximally prohibited.

The tendency to resist overweight lexical items also manifests itself in the formation of comparatives and superlatives in English. It is well known that polysyllabic adjectives, especially those of more than two syllables, tend to reject the *-er* and *-est* suffixes of comparison in favor of their periphrastic alternatives with *more* and *most*.

Note at this point that monosyllabic words account for over 80% of English text so that they are syllabically unmarked. On the other hand, words of more than two syllables account for only about 3% of English text so that they are syllabically marked. This may help explain at least in part why the first member of each pair below is perfectly acceptable while the second is not.

- (68) a. *singing/ringing vs. *lightninging*
 b. *clearer/dearer vs. ?eagerer/?properer*

Admittedly, the constraint on proximate repetition (Park 1977a, 1977b, 1982 and 1983) is partly responsible for the unnaturalness of the second member of each pair here. The first member of each pair in (68) is natural despite the proximate repetition of *-ing/-er* while the second is far less so, which is evidently due to the difference in the number of syllables between the two members. Note that the first member of each pair comprises only two syllables while the second comprises three. Note further that more than two syllables per (high-frequency) word may render the word in question overweight and thus marked, which is evidently why the proximate repetition

of *-ing/-er* is far less tolerable in the second member of each pair here than in the first. Given the perfect acceptability of the comparative *cleverer/severer*, however, there must be some other as yet unknown factor involved than just the proximate repetition and the number of syllables in the unnaturalness of the comparative *?eagerer* and *?properer*.

The following pairs of synonyms also testify to the tendency for words to keep from getting too polysyllabic and thus too overweight.

- (69) a. cultivable vs. cultivatable (<=cultivate + -able)
 b. educable vs. educatable (<=educate + -able)

The first member of each pair here is one syllable shorter, and thus that much lighter, than the second. Thus the first member of each pair above is closer to the canonical weight assigned to the lexical category than is the second, which is probably why the former is more frequently used than the latter.

More than five syllables per (low-frequency) word in *-able* may render the word in question syllabically a little too overweight. This may help explain why the second member of each pair below is definitely unnatural while the first is perfectly natural.

- (70) a. communicable vs. ?communicatable (<=communicate + -able)
 b. eradicable vs. ?eradicatable (<=eradicate + -able)
 c. eliminable vs. ?eliminatable (<=eliminate + -able)

Even five-syllable words in *-able* may not be really that tolerable, which is why the second member of each pair in (69) is not quite as natural as the first. Note also the interesting fact that the four-syllable *calculable* is perfectly acceptable while the five-syllable *calculatable* is not usually acceptable.

A similar desire to steer clear of overly long and weighty words is also apparently in evidence in data such as the following.

- (71) a. contiguity vs. *contiguo(u)sity (<=contiguous + -ity)
 b. continuity vs. *continuo(u)sity (<=continuous + -ity)
 c. ambiguity vs. *ambiguo(u)sity (<=ambiguous + -ity)
 (72) a. audacity vs. *audacio(u)sity (<=audacious + -ity)
 b. tenacity vs. *tenacio(u)sity (<=tenacious + -ity)
 c. frivolity vs. *frivolo(u)sity (<=frivolous + -ity)

Weight equilibrium for the lexical category may help explain why in the words below the word-final affricate, which was originally voiceless, is voiced in present-day English.

(73) knowledge/cabbage/cartridge/partridge

As already pointed out, the monosyllabic word is the norm in English so that words of more than one syllable exceed this norm. Thus there may arguably be pressure on words of more than one syllable to reduce their weight in one way or another so that they may better approximate the norm at least in weight, if not in the number of syllables. Given the fact that voicing has the effect of lowering the weight of the segment in question, we may speculate that the change of the original voiceless affricate into its voiced counterpart here is designed to reduce the weight of these disyllabic words to something closer to that of their more normal monosyllabic counterparts.

We may propose a similar account for the common British practice of voicing the word-final affricate in place names such as the following.

(74) Greenwich/Norwich

It is interesting that the affricate here does not get devoiced when these names refer to places in the New World, which apparently indicates that this voicing takes place over a considerable stretch of time as do all other significant sound changes.

Just as the word-final segment in *Greenwich* and *Norwich* alternates between voiced and voiceless affricates in British English, so does the final segment in *spinach* in most varieties of English. Needless to say, this voicing alternation in *spinach* must be identical in motivation to that in *Greenwich* and *Norwich*.

Now let us turn to the tendency to keep words from getting too underweight, as exemplified by the following words.

(75) a. *H*-bomb (<=**H* bomb<=*hydrogen* bomb)
 b. *Mideast* (<=**Mid* East<=*Middle* East)

The abbreviations *H* (for *hydrogen*) and *Mid* (for *Middle*) here are too lightweight to stand by themselves as independent orthographic words so that they get attached to what follows them. Note that this attachment does not

cause any serious weight problem since, among other things, the two original immediate constituents of either word here make up one single lexeme although they are spelled as two orthographic words.

The derivation of the following proper noun is amenable to a similar explanation.

(76) *Sinclair* (<=**Sin Clair*<=*Saint Clair*)

As we will see in 2.4, pre-noun forms tend to get considerably shorter than their underlying forms under the pressure of weight equilibrium. Thus the pre-noun (modifier) *Saint* in *Saint Clair* gets so reduced in weight that it is normally pronounced with the weakest vowel possible with the result that it usually is too lightweight to stand alone as an independent phonological or orthographic word. Thus it ends up getting attached to what follows, i.e. *Clair*, giving rise to *Sinclair*. Parenthetically, the deletion of /t/ from *Saint* in the derivation of *Sinclair* here is a case of consonant simplification identical, say, to the loss of /t/ in the derivation of *Christmas* from “*Christ + mas (s)*.” We may note in passing here that *Clair* may have added to the weight of the coda for *Saint*, thereby contributing to the weakening of its nucleus, in exactly the same way that *-mas(s)* has helped weaken the nucleus of *Christ-* in *Christmas*.

Note that contracted auxiliaries have to be suffixed to what immediately precedes them, as is shown below.

- (77) a. *I'm* coming. (<=**I m* coming. <=*I am* coming.)
 b. *She'll* come. (<=**She ll* come. <=*She will* come.)
 c. *It's* been okay. (<=**It s* been okay. <=*It has* been okay.)

Needless to say, the contracted auxiliaries here are too lightweight to stand by themselves as independent words so that they end up getting suffixed to what precedes them.

The writing of an acronym as one word is another example of the tendency to avoid overly underweight items under the pressure of weight equilibrium. Let us take the following acronyms for example.

- (78) a. AIDS (<=**A I D S*<=*acquired immune deficiency syndrome*)
 b. POWs (<=**P O Ws*<=*prisoners of war*)
 c. KO'd (<=**K O'd*<=*knocked out*)

d. CEOs (<=*C E Os<=chief executive officers)

The individual letters comprising each acronym here are too lightweight to serve as independent orthographic words so that they are put together into one written word despite their orthographically multi-word origin. Another contributory factor here is that the underlying source for an acronym is itself one lexical unit semantically despite orthographic indications to the contrary. Note in this connection that the acronyms here are formally more lexical than their underlying sources in that their inflectional suffixes are always word-final while this is not always the case with their sources.

2.4. Phrasal Weight

Phrasal weight is also more or less constant so that there is little or no weight variation between different manifestations of a phrase of the same category. Let us take the following paraphrase pair for example.

- (79) a. It *is a great honor* for me to do this.
 b. It *gives* me *great honor* to do this.

The verb *is* is more lightweight than the verb *gives* so that extra material like the indefinite article *a* in the verb phrase of (79a) above is a kind of ballast designed to make the verb phrase more or less equal in weight to that of (79b). Were this not the case, then weight equilibrium between the two verb phrases here could not be properly maintained. Needless to say, *honor* in (79) may be replaced with such words as *pleasure* and *delight* with exactly the same result.

A similar account applies to why we have the indefinite article in the verb phrase of the first member of the paraphrase pair below, but not in that of the second member.

- (80) a. The conference *was a success*.
 b. The conference *ended in success*.

Needless to say, *success* here may be replaced with *failure* with precisely the same result.

There was a time when the particles *in* and *out* had their respective variants *inne* and *oute*, as exemplified below.

- (81) a. He *is inne*.
 b. He *comes in*.
 (82) a. He *is oute*.
 b. He *goes out*.

The longer form of the particle in each pair here goes with the lighter verb *is* while the shorter form goes with the heavier verb *comes/goes*. The alternation between the two variants of the particle in each pair here is thus apparently designed to guarantee verb–phrase weight equilibrium between the two members of each pair above.

We have ample evidence that there is something like semantic weight equilibrium affecting English verb phrases. Let us consider the following pairs with specific reference to the difference in the semantic weight of the main (or head) verb between the first and second members of each pair.

- (83) a. Who *told* him?
 b. Who *told* him to go?
 (84) a. Let's *go*.
 b. Let's *go* fishing.
 (85) a. I *found* the book.
 b. I *found* the book interesting.
 (86) a. I will *keep* it.
 b. I will *keep* it a secret.
 (87) a. He finally *arrived*.
 b. He finally *arrived* in a jeep.
 (88) a. He *treated* me.
 b. He *treated* me with great respect. (=He showed me great respect.)

Note that the semantic weight of the main verb in each pair is in inverse ratio to the weight of its complement so that in (83), for one, *told* has heavier semantic content in the first member than in the second. Note further that thus the weight of the verb phrase remains more or less constant from one member of the pair to the other.

Needless to say, similar weight equilibrium is observable in each of the following pairs of sentences with copular verb phrases.

- (89) a. I am *anxious*.
 b. I am *anxious* to leave.

- (90) a. She is *good*.
 b. She is *good* at math.

The semantic weight of the (head) adjective in either pair here is in inverse ratio to the weight of its complement so that the verb phrase remains weight-wise more or less invariant from one member of the pair to the other.

Just as is the case with semantic weight, the head of a phrase normally has its phonological weight varied in such a way that it is in inverse ratio to the weight of its complement such as its modifier. Let us consider the following sentence pairs.

- (91) a. He's *good*.
 b. He's *very good*.
 (92) a. They're *teachers*.
 b. They're *good teachers*.

The head of the adjective phrase in (91) is normally heavier phonologically in the first sentence than in the second while the head of the noun phrase in (92) is normally heavier phonologically in the first sentence than in the second. This variation in phonological weight is evidently designed to guarantee weight equilibrium between the shorter phrase in the first member of each pair above and the longer one in the second member of the same pair.

As is the case with lexical weight, there is a tendency to keep a phrase from getting either overweight or underweight. Let us take a look at the following examples.

- (93) a. I *have to* help him
 b. I *have got to* help him.
 (94) a. I *may have to* help him.
 b. *I *may have got to* help him.

Although *have to* is in a kind of free variation with *have got to* in most instances, *may have to* is not in free variation with (the ungrammatical) **may have got to*. This is evidently because **may have got to* is overweight in that it is in excess of the canonical weight assigned to the auxiliary verb phrase. Needless to say, the use of other modals such as *will* instead of *may* in (94b) will result in an equally ungrammatical sentence.

The following pairs illustrate the tendency for phrases to keep from getting overly underweight.

- (95) a. *Smoke.*
 b. *Have a smoke.*
- (96) a. *Bathe.*
 b. *Take a bath.*
- (97) a. (?) *I photographed.*
 b. *I took a photograph.*
- (98) a. (?) *They copied.*
 b. *They made a copy.*
- (99) a. (?) *He goes.*
 b. *He is going.*

It appears that one-word phrases are generally too underweight for the phrase category, especially the verb phrase, which may be why the first member of each pair above is less natural than, and is thus less preferable to, the second member.

The desire to keep the verb phrase from getting too underweight may also explain at least in part why the first member of each sentence pair below is less natural than the second.

- (100) a. (?) *He leaves.*
 b. *He left.*
- (101) a. (?) *She arrives.*
 b. *She arrived.*

Since the past tense is heavier in semantic content than the present tense, the one-word verb phrase arguably is heavier, and thus more in compliance with the canonical weight of the verb phrase, in the second member of each pair here than in the first.

Our discussion here seems to throw light on the interesting fact that “The story *is interesting*” is acceptable while the logically plausible “The story *interests*” is not. It may very well be the case that “is interesting” is heavy enough to fill the predicate slot while “interests” is not. Needless to say, a similar account is applicable to the allowance of copular predicates ending in similar *-ing* adjectives (e.g. *demanding*), as opposed to the disallowance of those ending in the verbs from which the adjectives derive (e.g. *demand*).

Let us now turn our attention to weight equilibrium as it relates to noun phrases. Let us take a look at the following sentence pairs with special attention to the noun phrases in italics.

- (102) a. He is *a doctor*. (<=*He is *a Dr.*)
 b. He is *Dr. Smith*. (<=(?)He is *Doctor Smith.*)
- (103) a. Everest is *a majestic mountain*. (<=*Everest is *a majestic Mt.*)
 b. *Mt. Everest* is majestic. (<=**Mountain Everest* is majestic.)
- (104) a. He is *a saint*. (<=*He is *a St.*)
 b. He is *St. Paul*. (<=(?)He is *Saint Paul.*)

In each pair here, one and the same noun alternates between its longer and shorter forms, which occur in head-noun and pre-noun positions respectively. The noun in question here carries more information, and is thus much longer in form, as head noun than as pre-noun, which is as it should be (Park 1989). What is important to our discussion here is of course the fact that this positionally variant distribution in each pair here of one and the same noun helps maintain weight equilibrium between the two noun phrases in question. Note in this connection that we can similarly account for the alternation between *mister* and *Mr.* as well as between *mistress* and *Mrs.* Here, too, the longer forms are used as head nouns, and the shorter forms as pre-nouns.

The derivation of the indefinite article *a(n)* from *one* is also explainable as resulting from the pressure of weight equilibrium with *one* in pre-noun position optionally getting reduced to *a(n)*. A similar explanation also applies to the derivation of the definite article *the* from the demonstrative *that*. Needless to say, *that* optionally gets reduced to *the* in pre-noun position under the pressure of weight equilibrium, especially when its deictic meaning is weak.

Our discussion here throws light on the derivation of the word *another* from the pre-noun *one other*. The word *one* here has first arguably been reduced to *an* as part of the pre-noun *one other* under the pressure of weight equilibrium. Then the resulting pre-noun form, i.e. *an other*, has been further shortened to *another*, again under the pressure of weight equilibrium. In the second stage of the shortening here, the lightness of *an* may have lent a helping hand. As we have already shown in connection with our discussion of (75)–(78), lexical fragments, which are too light to stand alone

as independent words, tend to be attached to other words which are their immediate neighbors.

The derivation of *almost* (as in *almost all*) from the underlying premodifier of the form *all most* may be similarly accounted for. As with other premodifiers, the pressure of weight equilibrium may have forced *all most* to reduce its weight by getting a shorter form here, i.e. *almost*. A similar account is possible for pre-adverbial or pre-adjectival *al-* words such as *already* and *altogether*. Recall that we have dealt with other perspectives on the shortening process under discussion here in connection with (58), especially (58d), in 2.2.

Let us note at this point the difference in voicing between the word-final sibilant segments of the cognate pair below.

(105) *diverse*/dayvə:rs/ vs. *divers*/dayvə:rz/

The second adjective above may be used attributively only, i.e. in pre-noun position only, whereas the first adjective may be used either attributively or predicatively. There is thus apparently more pressure of weight equilibrium for reduction on *divers* than on *diverse* so that the word-final sibilant arguably gets voiced and thus lightened in the former, but not in the latter. Incidentally, *divers* is also orthographically one letter shorter than *diverse*, which may again be ascribed to the pressure of weight equilibrium.

Other pre-nouns than the denominal ones thus far considered also tend to involve reduction from their underlying sources, as can be seen from the following pairs of sentences.

- (106) a. Their welfare system is *from the cradle to the grave*.
 b. Theirs is a *cradle-to-grave* welfare system.

- (107) a. He is the economist *who was educated in the United States*.
 b. He is the *U. S. -educated* economist.

The longer form in the first member of each pair here is arguably the underlying source for the shorter pre-noun form in the second member of the same pair. The reduction involved in the derivation of the shorter pre-noun form here from its longer underlying source is of course also a product of the pressure of weight equilibrium.

Parenthetically, the use of *o'clock* is also apparently constrained by the principle of weight equilibrium as it applies to noun phrases. Note that ex-

pressions such as *6 o'clock* are acceptable while expressions such as *6:30 o'clock* are not. This is arguably because the former comply with the weight limit for the noun phrase while the latter do not. In other words, it is apparently the case that expressions like *6:30 o'clock* are too heavy to fill the canonical noun phrase slot while expressions like *6:00 o'clock* are not. An exactly identical account is applicable to the allowance of such expressions as *30-odd people* (= *thirty-odd people*), as opposed to the disallowance of *35-odd people* (= *thirty-five-odd people*).

Also of relevance to our discussion here is the fact that adjectives are generally more prominently stressed when they are the heads of adjective phrases (and thus potential adjective phrases) than when they are mere pre-nouns. Let us consider the following pair of sentences.

- (108) a. This book is *good*.
 b. This is a *good* book.

Normally, the adjective *good* here gets primary stress as the head of the adjective phrase in (108a) while it gets only secondary stress as pre-noun in (108b). As a result, *good* here is normally pronounced more prominently and longer in the former than in the latter. This suprasegmental variation of *good* is also due to the pressure of weight equilibrium.

Examples such as (102)–(108) strongly indicate that linguistic elements tend to assume shorter forms in pre-noun and other premodifier positions than elsewhere. As a general rule, premodifiers enter into close phrasal construction with whatever they premodify while their underlying allo-forms do not. Thus, other things being equal, the pressure of weight equilibrium for reduction is far greater on premodifiers than on their underlying allo-forms from which they derive. This pressure apparently throws light on why the pre-adjective *very* is adjectival in form instead of taking the usual adverbial form in *-ly*. Let us consider the following near-paraphrase pair.

- (109) a. He is *very* good.
 b. He is *verily* good.

The pre-adjective *very* (109a) is of adjectival form and thus exceptional in that pre-adjectives are generally manifested on the surface by *-ly* adverbs or adverbials. Thus *very* occurs here in a position where we would expect something like the adverbial *verily*, so to speak. We may argue here that

the hypothetical underlying form *verily*, as in (109b), has given way to the shorter (and thus lighter) *very* under the pressure of weight equilibrium.

The derivation of (110a) below from (110b) may also be explained in a similar manner as being due to the pressure of weight equilibrium on pre-modifiers for reduction.

- (110) a. He is *real* good.
 b. He is *really* good.

The following paraphrase pair contains another interesting instance of weight equilibrium at work.

- (111) a. I'll be *right* there.
 b. I'll be there *right away*.

The temporal adverbial here alternates between its shorter and longer forms, which occur inside and outside respectively of the core verb phrase, i.e. *be there*. The underlying adverbial *right away* has to be shortened to *right* inside the core verb phrase here so as not to render this core verb phrase overweight, while it is under no such pressure outside the core verb phrase.

Let us now take a look at weight equilibrium in action in prepositional phrases, as exemplified by the following sentence pair.

- (112) a. Here's *to John*.
 b. Here's *to you*.

In many people's normal pronunciation, the preposition *to* is weakly stressed in (112a) and strongly stressed in (112b) whereas its complement is strongly stressed in (112a), where it is a noun with inherent semantic weight, and weakly stressed in (112b), where it is a pronoun with little inherent semantic weight. The stress patterns here, i.e. 'weak+strong' for *to John* and 'strong+weak' for *to you*, are such that suprasegmental weight equilibrium is maintained between the two prepositional phrases in question.

The principle of weight equilibrium sheds light on the difference in usage between *said*, on the one hand, and *aforsaid* and *aforementioned*, on the other, as exemplified by the paraphrase pairs below.

- (113) a. *The said person* is responsible for the accident.
 b. **The said* is responsible for the accident.
- (114) a. *The aforesaid person* is responsible for the accident.
 b. *The aforesaid* is responsible for the accident.
- (115) a. *The aforementioned person* is responsible for the accident.
 b. *The aforementioned* is responsible for the accident.

We can see here that *said*, *aforesaid* and *aforementioned* are all light enough to serve as pre-nouns while *aforesaid* and *aforementioned*, but not *said*, are heavy enough to function as noun phrase substitutes. Note that *said*, which is the lightest of the three expressions here, is too light to qualify as a noun phrase substitute. What is clearly in evidence here is the tendency for noun phrases to keep from getting overly underweight.

Other things being equal, *aforesaid* and *aforementioned* are of course phonologically a bit lighter in pre-noun position than in head noun position. Thus *aforesaid*, for one, is slightly lighter in phonological weight in (114a) than in (114b), which is also in line with the principle of weight equilibrium.

The numeral *one* often has to replace the indefinite article *a(n)* in order to help avoid an excessively underweight phrasal slot. Let us take the following paraphrase pairs for example.

- (116) a. He was here for *one hour or two*.
 b. He was here for *an hour or two*.
- (117) a. He was here for *one or two hours*. (one=one hour/an hour)
 b. *He was here for *an or two hours*. (an=an hour/one hour)

We can see from this data that *one* is light enough to serve as a pre-noun and heavy enough to serve as a noun-phrase substitute. On the other hand, *a(n)* is light enough to serve as a pre-noun, but not heavy enough to serve as a noun phrase substitute.

The principle of weight equilibrium is instrumental in rationalizing the derivation of possessive pronouns, as opposed to possessive adjectives. Let us consider the following paraphrase pair.

- (118) a. This book is *your book*.
 b. This book is *yours*. (= *This book is *your*.)

Note that *your* alone is too light to fill the slot of the (subjective complement) noun phrase in (118b) here, which is equivalent to the slot occupied by *your book* in (118a). We resolve this problem by replacing *your* with *yours*, in which –s sort of stands for the head noun *book* in the noun phrase *your book*.

The principle of weight equilibrium also helps provide the rationale for the so-called expletives *there* and *it*, as exemplified below.

(119) There is somebody in the room. (<=(?)Somebody is in the room.)

(120) It is nice to see you. (<=(?)To see you is nice.)

The logical subjects here, i.e. the subjects of the parenthesized sentences, are informationally or structurally a little too heavy to fill the subject slot, which is canonically lightweight. To resolve this problem, we shunt the logical subjects out of the subject slot with the resulting vacuum being filled by the lightweight expletives *there* and *it*. The underlying subject slot here may not be left vacant because that would render the slot overly underweight. Thus we may argue that the expletives here are slot-fillers which are motivated by weight equilibrium and cataphoric to the logical subjects.

The so-called impersonal *it*, as exemplified below, can be similarly motivated.

(121) a. *It* is cold tonight. (*It*=The surroundings here/The weather here)

b. *It* is twelve noon. (*It*=The time now)

The impersonal *it* here is an exophoric pronoun referring to the general ambience which is too obvious to specifically refer to. Instead of a specific reference to the general ambience here, we use the lightweight filler *it* for the lightweight subject slot, again under the pressure of weight equilibrium. Note in this connection that not just expletives or impersonal pronouns but also all other pro-forms are slot-fillers motivated by weight equilibrium.

We have said earlier in this paper that the subject slot in English is canonically lightweight and thus not tolerant of heavy noun phrases while it is the other way around with the predicate slot. We have shown elsewhere, e.g. in Park 1981a, 1985 and 1989, that this is why sentences like “That he is a genius appears” is unacceptable while sentences like “It appears that he is a genius” is perfectly acceptable. We have also shown in the same pa-

pers that for the same reasons “Jack was sent a copy of the paper,” for example, is a far more natural passive transform for “We sent Jack a copy of the paper” than is “A copy of the letter was sent Jack.”

3. Surface Manifestation of Underlying Quality

Underlying quality in language tends to be maintained in surface form in one way or another with the result that underlyingly different linguistic entities tend to be differentially manifested on the surface (Park 1981a and 1989). Park (1989), for example, makes it abundantly clear that underlyingly weightier elements tend to be more clearly manifested on the surface than underlyingly weaker elements.

3.1. Phonological Manifestation

Expressive lengthening provides us with an interesting case of relevance to our discussion here. Let us consider the following word pairs.

- (122) a. *delicious* vs. ‘*de*licious’
 b. *delighted* vs. ‘*de*lighted’
 c. *long* vs. ‘lo-o-o-o-ong’

The second member of each pair here is emotively more emphatic than the first, which fact manifests itself in surface form in the use of a longer first-syllable nucleus for the second member than for the first.

The pronunciation of the suffix *-ful* affords another interesting example of surface manifestation of underlying quality. Let us compare the two sets of words given below.

- (123) a. *careful*/*beautiful*/*mournful* (/fəl/)
 b. *handful*/*mouthful*/*earful* (/fʌl/)

The suffix is slightly heavier semantically in (123b) than in (123a) in that the meaning of *full* is more concrete and distinct in (123b) than in (123a). It is also informationally heavier in (123b) than in (123a) in that the words in (123b) are nouns while those in (123a) are adjectives. For, other things being equal, nouns carry more information than adjectives, which is

often especially the case in examples like (123) above. For *handful*, for one, is usually a condensation of itself plus its deleted complement beginning with *of*. This difference in underlying weight manifests itself in the pronunciation of *-ful* with a heavier and more distinct vowel in (123b) than in (123a).

Note at this point that *farther* is referentially more concrete and thus heavier than *further* and that the heavier *farther* uses a heavier stem vowel than the lighter *further*. A similar phenomenon is observable with the phonological contrast between the deictic and less deictic uses of *that* and *there*. Identical in nature to this is the stress (and attendant segmental phonetic) difference between the locative *by*, as in “We live *by* the river,” and the non-locative *by*, as in “Man cannot live *by* bread alone.”

Each word pair below shows that one and the same prefix may be pronounced differently as a consequence of the underlying semantic difference between the two members of the pair.

- (124) a. *resign* vs. *re-sign*
 b. *proverb* vs. *pro-verb*
 c. *detect* vs. *defang*
 d. *predict* vs. *predate*

The prefix in each pair here is semantically heavier in the second member than in the first, which difference is reflected in its greater phonological prominence in the second member than in the first.

Note that the superficially identical stretch *I believe* is pronounced differently in the following pair of sentences.

- (125) a. *I believe* he is okay.
 b. He is okay, *I believe*.

I believe is semantically heavier in the first sentence of (125) above than in the second in that it expresses an assertion in the former while it is a mere comment in the latter. It is thus pronounced in less time and with less prominence in the second sentence than in the first. This is a surface reflection of the difference in underlying semantic weight between the two tokens of *I believe* here.

The law of conservation affords us insight into why the plural suffix *-s* is always voiceless in words such as the following.

- (126) a. months /mʌnθs/ or /mʌns/
 b. sevenths /sevɪθs/ or /sevns/

In the first pronunciation for either word here, the plural suffix follows a voiceless sound so that it is pronounced with the voiceless /s/. In the second pronunciation, the suffix follows a voiced sound and yet it remains voiceless without getting voiced to /z/ in apparent violation of a well-known rule of assimilation in English. This is evidently because the trace of the elided underlying voiceless /θ/ is still there, keeping the suffix from getting voiced. Thus viewed, the second pronunciation of either word here does not really violate the said rule of assimilation at all.

In connection with (126b), one may argue that the final sibilant segment in /sevns/ does not undergo voicing so as to make it distinct from /sevɪnz/ *sevens*, the plural of *seven*. Given the fact that /mʌns/ in (126a), which is a similar example, does not have to be distinguished from a word of the form */mʌnz/, plural or otherwise, however, we may reject this blocking argument as highly implausible and thus unconvincing.

The trace of an elided underlying segment may sometimes be completely obliterated, especially with the passage of time as well as with the frequency of occurrence. This may explain why the final sibilant segment of *Mrs.* is voiced today in spite of its voiceless origin in *Mistress*. It may also explain why the sibilant segment generally gets voiced in *asthma*, but not in *isthmus*.

The law of conservation offers a rare insight into an otherwise baffling phenomenon relating to the following example of neutralization from English, especially its prevalent American variety.

- (127) a. rider (= /raydər/ (= [ray dər]))
 b. writer (= /raytər/ or /raydər/ (= [raytər] or [raydər]))

Even when /t/ in *writer* is neutralized to /d/, the immediately preceding diphthong here is phonetically shorter than the one preceding /d/ in *rider*. Note that the /t/ neutralized to /d/ here is underlyingly voiceless and that the trace of this underlying voiceless quality causes the diphthong to be shorter here than before the underlyingly voiced /d/. The difference in the surface length of the diphthong in the two words here is thus due to the underlying difference in voicing between the two immediate post-diphthongal

consonants. Thus the trace of the underlying voiced/voiceless contrast shows up on the surface here.

The law of conservation is instrumental in explaining the vocalic laxing exemplified by the following word pairs.

- (128) a. bleed vs. bled ($\langle = *bleedd \langle = *bleded$)
 b. breed vs. bred ($\langle = *breedd \langle = *breded$)
 c. lead vs. led ($\langle = *leadd \langle *leaded$)

We may argue here that we have underlyingly a geminate ‘d’ cluster at the end of the second member of each pair here so that the laxing of the vowel from /iy/ to /e/ here is a case of vocalic laxing in immediate pre-cluster position. We are suggesting that the vocalic laxing here is identical in nature to that observable in such pairs as *keep/kept* and *deep/depth*.

The word-final consonant alternation exemplified by the following word pairs can be explained along similar lines.

- (129) a. send vs. sent ($\langle = *sendd \langle = *sended$)
 b. lend vs. lent ($\langle = *lendd \langle = *lended$)
 c. build vs. built ($\langle = *buildd \langle = ?builted$)

We may argue here that the original word-final /d/ combined with the underlying participial suffix $-(e)d$ to produce the word-final /t/ in the second member of each pair here. This is quite an interesting argument in view of the fact that the voiceless /t/, which is heavier than the voiced /d/, is well suited to stand for a combination of the two tokens of /d/ or of the stem-final /d/ and the participial $-ed$. We may note in passing here that this fortition phenomenon is observed immediately after the stem-penultimate segment /l/, /n/ or /m/ and that the stem-final /l/, /m/ or /n/ may have something to do with the word-final /t/ in participial forms like *felt*, *burnt* and *dreamt*.

The presence or absence of a major word-internal boundary often shows up on the surface in pronunciation differences between otherwise identical words. Let us take the following word pairs for example.

- (130) a. wife’s ($\langle = \text{wife} \#s$) vs. wives ($\langle = \text{wife} + s$)
 b. truth’s ($\langle = \text{truth} \#s$) vs. truths ($\langle = \text{truth} + s$)

In each pair here, the major boundary # separates the two constituent

morphemes of the first member far more clearly than the minor boundary + does the two constituent morphemes of the second member. Incidentally, the presence of the apostrophe in the first member of each pair, as opposed to its absence from the second, is orthographic testimony to the major and minor boundaries referred to here. On the basis of this boundary difference, we may say that the first member of each pair here comprises two clearly separate intralexical units while the second member is more or less a seamless word with the two intralexical units not quite as clearly separate. Thus we may argue that the major boundary in the first member of each pair above is strong enough to blunt the pressure of the suffix on the stem-final consonant to become more lightweight by getting voiced while the minor boundary in the second member of the same pair is not sufficiently strong to do so. Needless to say, we can similarly explain why we have /z/ in *resign/reserve* (= *re+sign/re+serve*), as opposed to /s/ in *re-sign/re-serve* (*re # sign/re # serve*).

The pronunciation difference between the adjective *tarry* and the verb *tarry* is due to the fact that underlyingly the former is *tar # -(r)y* while the latter is not thus analyzable. Note that the vowel of the noun *tar* remains basically unchanged in the adjective *tarry* because the major boundary here keeps the suffix from forcing the stem vowel to be laxated to /æ/.

The law of conservation is instrumental in explaining the apparent grapheme-to-phoneme gap illustrated by the following sets of words.

- (131) a. Chicago/Michigan/Cheyenne
 b. Illinois/Des Moines/Arkansas

All these words are underlyingly French in that they are all North American place names given by French colonists. Under French influence, the grapheme *ch* here is pronounced /ʃ/ while the word-final grapheme *s* is silent. This shows that the Frenchness of these words is preserved well into present-day English. English words of other foreign origins often tend to similarly retain their foreignness in distinctly non-English grapheme-to-phoneme correspondences, especially in ostentatiously educated speech. Let us take a look at the following examples.

- (132) a. Nike/psyche (Greek words with a non-silent final *e*)
 b. zeitgeist/Zwingli (German words with the word-initial /ts/)

It is interesting that some underlying ancient phonological segments are preserved in immediate pre-tonic position while they are weakened elsewhere, as can be seen from the two pairs of words below.

- (133) a. giant vs. *gigantic*
 b. devil vs. *diabolic*

The alternation between zero and /g/ in (133a) as well as between /v/ and /b/ in (133b) is apparently a joint product of the law of conservation and stress patterning.

It is interesting that what grammarians refer to as focus leaves a trace that gives rise to marked nuclear stress, as exemplified by the following sentences.

- (134) a. We like YOUR dress. (=We like your dress, not somebody else's.)
 b. WE like your dress. (=We, not somebody else, like your dress.)
 (135) a. Was he a FAMOUS singer? (=I know he was a singer, but was he a famous one?)
 b. Was HE a famous singer? (=I know she was a famous singer, but was he also a famous singer?)

Contrastive focus is on YOUR in (134a) and on WE in (134b) while interrogative focus is on FAMOUS in (135a) and on HE in (135b). These focused words get marked nuclear stress on account of the extra weight they gain from (the trace of) the focus they receive. A comparison of the sentences in (134) and (135) with their parenthesized paraphrases suggests that the focused words here contain not just their own weight but that of (the trace of) the material suppressed from the surface. The marked nuclear stress phenomenon under discussion here is yet another intriguing manifestation of underlying quality, which is often very elusive.

We can of course similarly account for marked nuclear stress due to other types of focus, as exemplified below.

- (136) a. They didn't EAT those apples. (=They didn't eat those apples; they only looked at them.)
 b. THEY didn't eat those apples. (=They didn't eat those apples; somebody else did.)
 (137) a. JOHN only read two pages. (=John read two pages; nobody

else did.)

- b. John only read TWO pages. (=John read two pages; he read no more.)

3.2. Orthographic Manifestation

Underlying quality often finds its surface manifestation in orthographic form. Let us begin our discussion here with the following pairs of cognates.

- (138) a. off vs. of
 b. too vs. to
 c. iff vs. if

Each pair here comprises two words of the same ultimate origin with the proviso that the first member is semantically heavier than the second. This difference in underlying semantic weight is reflected in surface form by the use word-finally of two tokens of one and the same letter in the first member of each pair here, as opposed to just one token of this letter in the second member. Parenthetically, this difference in underlying semantic weight is also reflected in the stronger pronunciation for the first member of each pair than for the second.

It is interesting that in each pair of homophones below the first member, which is a content word, is one letter longer than the second, which is a function word.

- (139) a. bee vs. be
 b. wee vs. we
 c. inn vs. in
 d. ass vs. as
 e. Ann vs. an

In our terms, the orthographic difference here is attributable to the fact that the contentive member of each pair above is underlyingly heavier in semantic content than its functional opposite number.

The two past participle forms of *bear*, as illustrated below, can also be accounted for in a similar manner.

- (140) a. I was *born* in Seoul in 1949.
 b. She has *borne* five children.

Note that *borne* is transitive while *born* is halfway between transitive and intransitive so that *borne* is slightly heavier semantically, and thus orthographically one letter longer, than *born*.

The following examples illustrate an interesting orthographic manifestation of underlying quality involving abbreviations.

- (141) a. ex. (<=example) vs. exx. (<=examples)
 b. p. (<=page) vs. pp. (<=pages)
 c. ms. (<=manuscript) vs. mss. (<=manuscripts)
 d. c. (<=copy) vs. cc. (<=copies)
 e. l. (<=line) vs. ll. (<=lines)
 f. f. (<=following page, etc.) vs. ff. (<=following pages, etc.)

Here the singular form ends in a single consonant letter while the plural form ends in a geminate cluster of this singular-final consonant letter. Since the plural is semantically heavier than the singular, what we see here is in perfect conformity with the law of conservation.

The difference in underlying semantic weight is often manifested in the use of upper- and lowercase letters, as illustrated in the following pairs of nouns.

- (142) a. *god* vs. *God*
 b. *lord* vs. *Lord*
 c. *scripture* vs. *Scripture*
 d. *fate* vs. *Fate*

The first member of each pair here normally refers to a lesser being or entity than does the second member. For example *god* refers to a mere deity while *God* refers to the Creator of the universe. Thus the second member of each pair above is semantically heavier than the first, which translates into the capitalized initial letter of the second member, as opposed to the use of a lowercase letter for the initial of the first member. Note in this connection the general practice of capitalizing the initial letter of the third-person masculine singular pronoun when *God* is referred to. Note also that the great playwright William Shakespeare is often referred to as *the Bard* with a capital *B* although a lowercase *b* here is not uncommon.

Our next example has to do with the common practice of beginning with a capital letter a common noun or adjective used as a term of address, as il-

illustrated by the following sentences.

- (143) a. Where were you, Captain?
 b. Hi, *Beautiful!*

Note that a term of address has the function of singling out one particular person in the immediate environment as the recipient of the linguistic message to be communicated. As such, the common noun or adjective in question carries more information than it does elsewhere so that this added information often translates into the (optional) capitalization of the initial letter of the noun or the adjective as the case may be.

We may say that the vocative in (143a) is a common noun converted into a proper noun while its counterpart in (143b) is a common adjective converted into a proper adjective. As in these two cases, a common noun or adjective gets an initial capital letter when it is converted into a proper noun or adjective. Needless to say, this is because a proper noun or adjective is (referentially unique and thus) informationally heavier than its common noun or adjective counterpart.

Of relevance to our discussion here is the common practice of indicating special emphasis by capitalization, underlining, italicization or gothicization, which is a surface manifestation of the extra semantic weight the writer intends to convey to the reader. As regards the practice of capitalizing initial letters in block language, note that prepositions usually do not allow such initial capitalization with the exception of long prepositions, usually those with more than four or five letters. Note here that longer prepositions tend to be heavier in semantic content than shorter ones so that the allowance of initial capitalization for longer prepositions, as opposed to its disallowance for shorter ones, is a consequence of the difference in underlying semantic weight between the prepositions in question.

The comma is often used as a surface marker of underlying differences, as can be seen from the following pair of coordinate phrases.

- (144) a. X, and Y or Z (=either X + Y or X + Z)
 b. X and Y, or Z (=either X + Y or Z)

Needless to say, *X and Y or Z* without a comma properly positioned would be two ways ambiguous between (144a) and (144b).

Let us compare the following two sentences with special attention on the

boundary between the respective constituent clauses.

- (145) a. I'll tell you later *if I can spare the time*.
 b. I'll tell you later(,) *if I can spare the time*.

The *if*-clause is intended to be the direct object of *tell* in the first sentence of (145) while it is intended to be a conditional adverbial in the second. Thus the subordinate clause is tied more closely to the main clause in the first sentence than in the second, which difference manifests itself in the absence of the comma from the former, as opposed to its optional presence in the latter.

Let us now consider the following sentence pairs.

- (146) a. I'm not going.
 b. I'm not going, *period*.
 (147) a. He's not acceptable.
 b. He's not acceptable, *full stop*.

The second member of each pair here is far more emphatic than the first member. For the finality/definiteness that is ordinarily indicated by the period or full stop is far more emphatically expressed by its spelled-out or verbal version, i.e. *period* or *full stop*.

The dieresis used in words such as the following provides us with yet another case of underlyingly determined surface differentiation of two otherwise indistinguishable orthographic segments.

- (148) a. *cooperate/co-operate/coöperate*
 b. *reexamine/re-examine/reëxamine*

Note that the doubling of the letter *o* or *e* here is problematic at least on two counts. Firstly, the double *o* and *e* here may be confused with the digraphs *oo* and *ee* respectively. Secondly, proximate repetition of any linguistic element, orthographic or otherwise, is a burden on the encoder as well as the decoder so that it tends to be severely constrained in natural language (Park (1977a, 1977b, 1982, 1983)). To resolve this problem, we often have recourse to a hyphen between the two tokens of the same letter in question here or to a dieresis on the second of the two tokens to set them clearly apart. The dieresis is put on the letter belonging to the stem here, not on the one belonging to the prefix. Thus we can say that the heavier of

the two letters in question here gets the dieresis, which is a faithful surface reflection of the difference in underlying weight between the two letters in question.

3.3. Grammatical Manifestation

Let us begin our discussion here with the following paraphrase pairs.

(149) a. L.A. is worth *visiting at least once*.

b. L.A. is worth *at least one visit*.

(150) a. L.A. is worth *visiting many times*.

b. L.A. is worth *many visits*.

Assuming that the first member of each pair here is the underlying source for the second, we can see that the number distinction in the noun phrases here may be a relatively surface phenomenon which originates in the number distinction in their underlying verb–phrase sources.

The following examples also show that the surface number distinction often originates in a deeper–level number distinction.

(151) a. a *branch* director (<=a director in charge of a *branch*)

b. a *branches* director (<=a director in charge of *branches*)

(152) a. a *career* girl (<=a girl pursuing a *career*)

b. a *careers* girl (<=a girl dealing with *careers*)

Assuming that the parenthesized paraphrases are possible underlying sources for the noun phrases here, we can argue that the surface distinction between singular and plural for the premodifying nouns here is determined in their underlying sources.

The number distinction between the postmodified nouns in the following noun phrases is also apparently determined at an underlying level.

(153) a. the *case* of Germany and France

b. the *cases* of Germany and France (<=the case of Germany and the case of France)

Note that underlyingly France and Germany are treated as one single unit of consideration in (153a) while they are treated as two distinct units in (153b). Thus the choice between the singular *case* and the plural *cases*

here is made on the basis of this deeper-level distinction.

It is well known that there is usually no surface manifestation of the underlying singular/plural distinction for the second-person pronoun in English. However, many people do use the following plural forms for the second person.

- (154) you people/you guys/you fellows/you boys/you gals

Thus in the speech of those who use these forms at least, the underlying singular/plural distinction for the second-person pronoun in English does manifest itself on the surface.

Similar in nature to the use of geminate consonants for plurality, as illustrated in (141) above, is the use of conjunctive reduplications for repetitive or emphatic meaning, as illustrated in (155)–(161) below.

- (155) a. for *ever*
 b. for *ever* and *ever*
- (156) a. for *years*
 b. for *years* and *years* and *years*
- (157) a. *again*
 b. *again* and *again*
- (158) a. She *cried*.
 b. She *cried* and *cried* and *cried* and *cried*.
- (159) a. We saw *dogs*.
 b. We saw *dogs* and *dogs* and *dogs*.
- (160) a. She is *very* good.
 b. She is *very, very* good.
- (161) a. Prices kept going *up*.
 b. Prices kept going *up, up, up*.

The second member of each pair above is more emphatic than the first, and the varied emphasis is expressed on the surface by varying degrees of conjunctive repetition. As compared to their non-reduplicative counterparts, in fact, all reduplicatives are kind of emphatic with the emphasis being expressed by formal repetition.

Note in this connection that the plural suffix often takes the place of conjunctive reduplication in expressing repetitive emphasis. This usage of the plural suffix is illustrated by the following examples.

- (162) a. the mumps/the chicken pox/the jitters/the measles
 b. the rains/the snows/the woods/the sands/the rapids/the waters

These plural expressions refer to conditions involving numerous repetitions of what is denoted by their singular counterparts or to a stretch or expanse comprising such repetitions. We are suggesting here that *rains* in (162b), for one, derives from an underlying structure of the form *rain and rain and rain and rain and rain*...

The present participial form also arguably takes the place of conjunctive reduplication, as can be seen from the following sentence pair.

- (163) a. I *tapped* on the table.
 b. I *was tapping* on the table.
 (164) a. I saw him *shoot* at the crowd.
 b. I saw him *shooting* at the crowd.

(163a) and (163b) imply one single tap and a series of taps respectively while (164a) and (164b) denote one single shot and a series of shots respectively. Here, too, a longer linguistic form expresses a longer real-life stretch of activity than does its shorter counterpart, which complies with the law of conservation.

Additional length, whether due to the plural marker or reduplication or otherwise, generally stems from additional weight of meaning or emphasis. Note that *not any more than*, for one, is more emphatic and thus longer than *no more than* so that (165a) below is more emphatic than (165b).

- (165) a. I am *not any more* a genius *than* he is a fool.
 b. I am *no more* a genius *than* he is a fool.

(166a) below is similarly more emphatic than (166b) so that the former is longer than the latter.

- (166) a. John is much *too mean for him to* lend you money.
 b. John is much *too mean to* lend you money.

Additional length of an expression often takes the form of an extra article, usually the definite article, as can be seen from the following pairs.

- (167) a. He spoke in *defence* of justice.
 b. He gave evidence against *the defence* in the Dreyfus case.

- (168) a. *Resistance* to the law was negligible.
 b. *The resistance* refused to negotiate.
- (169) a. *Opposition* to the new leader was extremely fierce.
 b. *The opposition* consisted of my father and a few of his friends.
- (170) a. He was in *prison*.
 b. He was in *the prison*.
- (171) a. He left at *sunrise*.
 b. He was watching *the sunrise*.
- (172) a. *Summer* is here.
 b. He spent *the summer* here.
- (173) a. The prime minister went on *television*.
 b. My brother turned off *the televisor*.

In each pair here, one and the same noun is preceded by the definite article in the second member, but not in the first. The noun without the article here has an abstract (zero-dimensional) meaning while the one with the article has a concrete (dimensional) meaning. Given the fact that an abstract meaning is generally lighter than its concrete counterpart, we may argue here that the addition of the article to the noun in question in the second member of each pair reflects the fact that this token of the noun here is semantically heavier than that of the same noun (minus the article) in the first member of the same pair. Parenthetically, a similar explanation apparently applies to the minimal pairs *in front of vs. in the front of*, *in back of vs. in the back of*, and *on top of vs. on the top of*.

In examples such as the following pair of expressions also, the definite article apparently serves to indicate greater semantic weight than does the zero article.

- (174) a. New Year (=New Year's Day)
 b. *the* New Year (=the first few weeks of a year)

(174a) refers to a shorter period of time than does (174b) so that here again difference in semantic weight translates into difference in surface length in a rather straightforward manner.

Also resulting from a difference in underlying semantic weight are the two past participle forms of the verb *get*, i.e. *got* and *gotten*, whose distribution is illustrated by the following sentence pairs.

- (175) a. I *have got* a ticket. (*have got*=*have*)
 b. I *have gotten* a ticket. (*have gotten*=*have acquired*)
 (176) a. I *have got to* do more work. (*have go to*=*must*)
 b. I *have gotten to* do more work. (*have gotten to*=*have come to*)

As the parenthesized paraphrases here show, *gotten* is semantically heavier than *got* in either pair above. Here again this difference in underlying semantic weight between the two participle forms translates quite straightforwardly into a difference in surface length. Likewise, *drunken* is semantically heavier and thus formally longer than *drunk*, as in *drunken driving* vs. *drunk driving*.

Let us consider the difference between the following two sentences with regard to the sentence-initial adverbial clause.

- (177) a. *Whether right or wrong*, he always takes sides with her.
 b. *Whether rightly or wrongly*, he always takes sides with her.

The sentence-initial adverbial clause in the first sentence here normally qualifies the subject of the superordinate clause although it can also qualify the entire superordinate clause. On the other hand, the sentence-initial adverbial clause in the second sentence always qualifies the superordinate clause. Thus, the adverbial here is arguably more lightweight in the first sentence than in the second. This is reflected in the absence of the *-ly* suffix from the adverbial of (177a), as opposed to its presence in the adverbial of (177b), which is in line with our law of conservation. Incidentally, we can see from examples like (177) that derivatives (e. g. *wrongly*) are generally longer in form than their sources (e.g. *wrong*) because the former bears more meaning than the latter. This phenomenon is discussed in detail in Park (1980, 1981b and 1989).

Let us compare “*Jack and Jill saw each other*” and “*Jack saw Jill and Jill saw Jack*.” The first sentence normally denotes that “*Jack saw Jill and, at the same time, Jill saw Jack*,” while the second sentence normally denotes that “*Jack saw Jill and then Jill saw Jack*.” Thus the first sentence implies a shorter span of time (or activity) than does the second, which difference is reflected in the fact that the first sentence is physically shorter and more compact than the second. This again is in line with our law of conservation in that the size of the surface form of an expression is in direct proportion

to its underlying semantic weight.

Let us take a look at the following paraphrase pairs with specific reference to the contractability of the modal *will*.

- (178) a. He *will* refuse. (*will* of insistence)
 b. *He'll refuse.
 (179) a. He *will* help us. (*will* of simple future)
 b. He'll help us.

The *will* of insistence here is semantically much heavier than the *will* of simple future with the result that the latter, but not the former, allows reduction to *-'ll*. Thus the surface contractability of *will* here is a function of its underlying semantic weight.

Content words are semantically heavier than function words so that the former are generally pronounced more prominently than the latter. Thus when one and the same word may be used as either a content word or a function word, then it is pronounced more distinctly as a content word than as a function word. For example, *the* is phonologically more prominent when it has the contentive meaning of "the most remarkable" than when it is a plain definite article. Other frequently cited cases in point include *there* (which can be either a contentive place adverbial or a functional expletive adverbial) and *that* (which can be either a contentive demonstrative or a functional subordinator).

Another consequence of the difference in semantic weight between content and function words is that the latter, but not usually the former, allow contraction. Let us consider the following paraphrase pairs.

- (180) a. I *have* been there already.
 b. I've been there already.
 (181) a. We *have not* seen her yet.
 b. We've not seen her yet./We haven't seen her yet.
 (182) a. The *haves* must remember the *have-nots*.
 b. *The'ves must remember the've-nots./*The'ves must remember the haven'ts./*The'ves must remember the have-nots./*The *haves* must remember the've-nots./
 *The *haves* must remember the *haven'ts*.

We can see here that the functional (=auxiliary) *have* may be contracted

to *-'ve* while the contentive *haves* may not be contracted to *-'ves*. We can also see that the functional *have not* allows contraction to either *-'ve not* or *haven't* while the contentive *have-nots* does not allow contraction to either *-'ve-not* or *haven'ts*. Needless to say, the contractability of *have (not)* here is also a function of its underlying semantic weight.

Note at this point that the lightweight verb *have*, be it auxiliary or main verb, may contract to *-'ve* with the proviso that it is more likely to do so as an auxiliary than as a main verb. It goes without saying that this is again due to the difference in underlying semantic weight between the contentive and functional uses of one and the same verb *have*.

Whether functional or contentive, the past *had* is less amenable to contraction than the present *have* because *had* is underlyingly heavier than *have* in that it originates in *have+past tense*. As is the case with *had* here, auxiliary verbs in general gain considerable weight when extra material is added thereto, as can be seen from the following pairs.

- (183) a. *have* = > *-'ve*
 b. *have not* = > *haven't* = > **-'ven't*
- (184) a. *is* = > *-'s*
 b. *is not* = > *isn't* = > **-'sn't*
- (185) a. *would* = > *-'d*
 b. *would not* = > *wouldn't* = > **-'dn't*

We can observe here that auxiliaries, which may ordinarily be reduced to their contracted forms, may not be thus shortened when they are already suffixed with the negative contraction *-n't*. Suffixed with the negative contraction, an auxiliary gets much heavier than it would otherwise be. Thus an auxiliary suffixed with the negative contraction is kind of too heavy to be attached to the immediately preceding (pro)noun because the resulting word would be too overweight. This is apparently in large measure why double contraction is not allowed in cases like the second member of each pair above. Again the contractability of the auxiliary here may largely be determined by the underlying weight of the word of which it is a part. The constraint on proximate repetition (Park (1982)) is also apparently a variable of relevance here just as it is in the case of (67b) in 2.3.

Let us now turn to the apparent anomaly of prepositions or subordinate conjunctions governing adjectives, as exemplified by the following sentences.

- (186) a. This is *far from (being) complete*.
 b. That was *close to (being) miraculous*.
 (187) a. Work hard *while (you are) young*.
 b. We will do so *if (it is) possible*.

The parenthesized portions are optional and often better left out. Part of the reason for this is of course to be found in the semantic lightness of the elements concerned. Thus the prepositions in (186) and the subordinate conjunctions in (187) may end up (superficially) governing adjectives, which is apparently anomalous. We can account for this odd phenomenon by saying that the deleted parts have their (invisible) traces left in their place so that the prepositions and conjunctions here govern these traces plus whatever follows.

Incidentally, a (de-conjunctive) preposition may end up (superficially) governing another (de-conjunctive) preposition in examples of the following sort.

- (188) We played tennis *until (it was) after (it was) dark*.

The two parenthesized portions here are light enough to be candidates for deletion and that at least one of them has to be deleted under the pressure of the constraint on proximate repetition (Park (1977a, 1977b, 1982, 1983)).

We can see from (186)–(188) that the verb *be* is so light that it often gets deleted. Not all instances of *be* may be so freely deleted, however, as can be seen from the following paraphrase pairs.

- (189) a. He seemed *to be* asleep.
 b. He seemed asleep.
 (190) a. He seemed *to be* upstairs.
 b. *He seemed upstairs.

The verb *be* here is a mere copula in (189a) while it has the contentive meaning of 'exist' in (190a), so that it is semantically heavier in the latter than in the former. Note that the lighter *be* here may be deleted from (189a) in the derivation of (189b) while the heavier *be* may not be deleted from (190a) so that (190b) is ungrammatical. This is yet another piece of evidence that underlyingly heavier elements tend to be better preserved on

the surface than lighter ones.

As we have seen repeatedly in this chapter, traces of deleted elements often exert exactly the kind of influence that the elements themselves would if they were not deleted. Let us take a look at the following paraphrase pair.

(191) a. I requested that he *should not be* included.

b. I requested that he *not be* included.

In (191a), *should* is kind of presupposed semantically by *request* so that it is largely redundant and thus light in information content, which is why it is often deleted, as in the derivation of (191b) from (191a). This is yet another example of a linguistic element getting deleted because of its light underlying weight.

Note that the deletion of *should* from (191a) here results in (191b) with the stretch *not be*. This stretch is apparently anomalous in two ways in that *be* does not seem to agree in number with *he* and in that *be* should normally precede *not*. These apparent anomalies are resolved as soon as we remember that the trace of *should* is very much in place in front of the stretch in question, exerting the influence of *should* itself as if it were not deleted.

We may speculate here that the trace of *should* is piggybacking on *be* in (191b) above so that *be* gets heavier here than would otherwise be the case and is thus in fact slightly heavier than *not*. If this is correct, then we can argue that it is only natural for the heavier *be* to follow the lighter *not* here in accordance with the light-to-heavy progression of communicative dynamism in English (Quirk et al. (1356–7)).

We may just as plausibly speculate that the trace of *should* may sometimes be piggybacking on *not*, rather than on *be*, when *should* is deleted from a sentence like (191a). In that case *not* should get heavier than *be* so that *be* should precede *not* in compliance with the normal communicative dynamism of English. This is indeed what happens in “I requested that he *be not* included,” which is a perfectly grammatical equivalent of (191b).

The utterance-final occurrence of *not* in Speaker B’s response in the following dialogue can be accounted for by having recourse to the concept of trace under discussion throughout this paper.

(192) Speaker A: Is Bob feeling better?

Speaker B: I'm aftaid *not*.

Note that *not* here is short for *he is not feeling better* and thus contains the trace of this suppressed clause. As a result, *not* here is far heavier than its normal token, as in *I'm not afraid*, so that its utterance-final occurrence in Speaker B's response above is in line with the normal communicative dynamism of English.

The following paraphrase set also testifies to the very real power of the trace left behind by a deleted element.

- (193) a. May God bless you.
 b. God bless you.
 c. God bless.
 d. Bless you.

Assuming that (193a) is the underlying source for the rest of the sentences here, we can say that the trace of *May* is very much in evidence in (193b) and (193c) so that the apparent number discord between *God* and *bless* is not really a discord. The trace of *May* is also evidently operative in the optative interpretation of (193b), (193c) and (193d). When *you* is also dropped from (193a), as in (193c), then we end up with the apparent anomaly of the transitive *bless* functioning here as an intransitive. However, the trace of *you* is also very much alive with its referent right there in the immediate speech context so that this apparent anomaly is not really an anomaly.

Parenthetically, (193a) may not normally be reduced to the one-word sentence of the form *Bless*. This may be designed to keep the sentence (or, more generally a syntactic phrase) from getting too underweight, i.e. from not exceeding the lower limits of its weight. Of relevance to our remarks here is our discussion of (95)–(101) in 2.4.

Let us now turn to the fact that *want to* may be reduced to /wanə/ in the first sentence below, but not in the second.

- (194) a. I *want to* work here.
 b. He is the man I *want to* work here. (< =He is the man. + I *want the man to* work here.)

In (194b), the trace of *the man* separates *want* and *to* and thus keeps *want*

to from getting collapsed into /wanə/. On the other hand, there is no such trace in (194a) so that there is no obstacle to the reduction of *want to* to /wanə/ here.

The definite article often needs to be taken as co-occurring with the trace of a common noun deleted from a complex proper noun. Let us take a look at the following paraphrase pairs.

- (195) a. the Sea of Galilee
 b. the Galilee
 (196) a. the House of Commons
 b. the Commons

The first member of each pair here, which is a complex proper noun, is arguably the underlying source for the second member. Note that in the first member of each pair the definite article goes with the head noun, i.e. *Sea* in (195a) and *House* in (196a). In the second member, on the other hand, the definite article is superficially transferred to the position immediately before the other noun, which is not the underlying head noun. For, in the second member, we no longer have the underlying head noun to which to attach the article. We can say that the deleted underlying head noun is represented in the second member of each pair above by its invisible trace afloat somewhere in the remainder of the expression, say, immediately after the definite article.

Only by positing this intangible trace can we satisfactorily resolve the surface anomaly of the definite article used with a proper noun/adjective in the second member of each pair below.

- (197) a. the Pacific Ocean
 b. the Pacific
 (198) a. the Ritz Hotel
 b. the Ritz

Note that in (195b), (197b) and (198b), the definite article goes not really with the proper noun/adjective but with the trace of the underlying head noun. Needless to say, the definite article in (196b) also goes not with *Commons* but with the trace of the deleted underlying head noun *House*.

We can similarly resolve the surface anomaly of the plural suffix on the proper adjective in the second member of each paraphrase pair below only

by positing a trace for the common noun deleted from the first member of the same pair.

- (199) a. the Rocky Mountains
 b. the Rockies
 (200) a. the Philippine Islands
 b. the Philippines

We are suggesting here that the traces of the deleted common nouns *mountain* and *island* here are invisibly piggybacking on (probably the tail ends of) *Rocky* and *Philippine* respectively. Thus the plural suffix in the second member of each pair above is arguably attached to the piggybacking trace of the common noun in question, not really to the proper adjective.

We may argue here that *linguistics* and other names of sciences in *-ics* end in the same plural suffix as that in *the Rockies* and *the Philippines*. We may suggest the following as possible underlying sources for *linguistics*, for example.

- (201) a. things linguistic
 b. linguistic matters

The nouns *thing* and *matter*, which are lightweight enough to be deletion-prone, are arguably deleted from (201a) and (201b). Then the stranded *-s* winds up piggybacking on the trace of *thing* or *matter*, which in turn is piggybacking on (the tail end of) the adjective *linguistic*, thereby giving rise to *linguistics*. We can thus resolve the surface anomaly of the plural suffix being attached to an adjective in such words as *linguistics*.

Derivationally quite similar to names of sciences in *-ics* are such words as *perishables* and *valuables*, which may be traced to (202a) and (202b) below respectively.

- (202) a. perishable things
 b. valuable things

Needless to say, the lightweight *thing* is first deleted here with its trace piggybacking on the immediately preceding adjectives, i.e. *perishable* and *valuable*. Then the plural suffix *-s* winds up getting a ride on the trace of *thing* already piggybacking on the adjectives here. Parenthetically, Park (1989) offers an interesting perspective as to why *things* is premodified in

examples like (202) while it is postmodified in examples like (201a)

The apparent anomaly of a singular noun ending in the plural suffix, as in the examples below, is also resolvable by having recourse to the underlying semantic structure of the noun in question.

- (203) a. *a sawbones* (< = a person who saws bones)
 b. *a butterfingers* (< = a person who has buttery fingers)
 c. *a thicklips* (< = a person who has thick lips)

Assuming that the underlying structure of each noun here is as given in the parentheses, we can argue that *person who (has)* is so lightweight that it gets deleted with its trace sort of in place between the indefinite article and the plural noun. Thus the plural suffix here does not have to clash at least underlyingly with the singular-denoting indefinite article. Needless to say, we get the indefinite article *a* here because the reference is to one person only while we get the plural suffix here because the thing involved comprises more than one unit. Thus both the indefinite article and the plural suffix in each noun phrase of (203) above are legitimate surface manifestations of integral parts of the underlying semantic structure of the noun in question.

It is worthy of note at this point that the semantic interpretation of bahuvrihi compounds must have crucial recourse to our concept of trace. The bahuvrihi compound *birdbrain*, for one, arguably derives from an underlying source like *a person who has a birdbrain*. We may assume that *a person who has* here is semantically so lightweight that it gets deleted in the deep-to-surface mapping process just as it does in the derivation of (203b) and (203c).

Our discussion here also appears to throw light on the phenomenon of reclassification, as exemplified by sentences such as the following.

- (204) a. He aspires to be a future *Shakespeare*.
 b. *Two beers*, please.

Shakespeare and *beer*, which are both originally non-count nouns, are used here as count nouns. We can reasonably assume that *a Shakespeare* in (204a) derives from something like *a person like (the) Shakespeare* while *two beers* in (204b) derives from something like *beer for two people, two glasses/bottles/cans of beer* or the like. Given the underlying sources for *a Shake-*

speare and *two beers* suggested here as well as our discussion of (199)–(203), the reclassification of the two nouns here from non-count to count offers us still another piece of evidence that underlying quality tends to be preserved in deep-to-surface mapping in human language.

Let us now turn our attention to the surface conservation of underlying order and distance between linguistic elements (Park 1985). It is interesting that nominalizations tend to maximally retain their deeper-level word order, as can be seen from the following paraphrase pairs.

(205) a. JFK was assassinated by Oswald.

b. JFK's assassination by Oswald

(206) a. Oswald assassinated JFK.

b. Oswald's assassination of JFK

Note the same word order in both the underlying clause and its nominalization, which clearly attests to the tendency for surface structure to keep underlying word order as intact as possible.

The following paraphrase pairs afford us another good example of word-order preservation.

(207) a. The kids came *to possess* the pills.

b. The kids came *into possession of* the pills.

(208) a. The pills came *to be possessed by* the kids.

b. The pills came *into the possession of* the kids.

Assuming that the first member of each pair here is the underlying source for the second member, we can see that the word order does not change from underlying to surface structure. We may further observe that the definite article in *came into the possession* in (208b) is equivalent to, and is arguably the trace of, '*be + -ed*' of *came to be possessed* in (208a).

Note in this connection that the passive equivalent (208b) is longer than the active equivalent (207b) just as the passive underlying source for the former is longer than the active underlying source for the latter. This goes to show that relative underlying length also tends to be preserved in surface structure.

The contrast between the 'active' *came into possession of* and the 'passive' *came into the possession of* may be somewhat productive in view of exactly the same contrast to be found in such pairs as *in view of/in the view of* and

in charge of/in the charge of, in which *the* is equivalent to the passive 'be + -ed by'. Recall at this point the comment we have made in connection with (167)–(174) that the (definite) article is often expressive of heavier meaning than is zero article.

Our examples (205)–(208) also show that the relative distance between major constituents in underlying structure tends to be preserved in surface structure. For good measure, we may consider the following paraphrase pairs.

- (209) a. The economist was educated in America.
 (economist >> educated >> America)
 b. the American-educated economist
 (America(n) << educated << economist)
- (210) a. The boy looks sick. (boy >> look(s) >> sick)
 b. the sick-looking boy (sick << look(ing) << boy)
- (211) a. The sound of the bell grew faint. (sound >> grew >> faint)
 b. Faint grew the sound of the bell. (faint << grew << sound)

The parenthesized material is designed to show that word order is reversed from the first member of each pair here to the second. Except for this reversal of word order, the relative distance between the major elements here remains the same in both the underlying source and its surface transform.

References

- Bronstein, A. J. (1960) *The Pronunciation of American English*, Prentice-Hall, Englewood Cliffs, N. J.
- Chomsky, Noam and Morris Halle (1968) *The Sound Pattern of English*, Harper and Row, New York.
- Gimson, A. C. (1980) *An Introduction to the Pronunciation of English*, Third Edition, Edward Arnold, London.
- Jespersen, Otto (1909–1949) *A Modern English Grammar on Historical Principles*, especially I and VI, Munksgaard, Copenhagen.
- Kreidler, Charles W. (1989) *The Pronunciation of English*, Basil Blackwell, Oxford.
- Park, Nahm-Sheik (1977a) 'Variables in Sentential Readability with Special Reference to Korean Learners of EFL/ESL,' Georgetown University Ph. D. Dissertation.

- Park, Nahm-Sheik (1977b) 'Structural Density as a Comprehensibility Variable,' *Language Research* 13.2, 95-103, Seoul National University, Seoul (In Korean).
- Park, Nahm-Sheik (1978) 'Contiguity as a Comprehensibility Variable,' *Language* 2.2, 65-74, The Linguistic Society of Korea, Seoul (In Korean).
- Park, Nahm-Sheik (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).
- Park, Nahm-Sheik (1981a) 'Weight as a Linguistic Variable,' an unpublished paper read at the Annual Spring Meeting of the Linguistic Society of Korea.
- Park, Nahm-Sheik (1981b) 'Distance and Length in Language with Special Reference to English,' an unpublished paper read at a seminar co-sponsored by the East-West Center and the University of Hawaii Department of Linguistics.
- Park, Nahm-Sheik (1982) 'Constraint on Proximate Repetition with Special Reference to English,' *Language Research* 18.1, 217-231, Seoul National University, Seoul.
- Park, Nahm-Sheik (1983) 'Proximity and Reduction,' *Language Research* 19.1, 75-102, Seoul National University, Seoul.
- Park, Nahm-Sheik (1984) 'The Constraint on Proximate Repetition and Morpho-phonological Explanation in English,' *Language Research* 20.4, 367-415, Seoul National University, Seoul.
- Park, Nahm-Sheik (1985) 'Natural Word Order in English,' *English Studies* 9, 11-32, Seoul National University, Seoul (In Korean).
- Park, Nahm-Sheik (1989) 'Weight as a Linguistic Variable with Reference to English,' *Language Research* 25.4, 711-894, Seoul National University, Seoul.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik (1985) *A Comprehensive Grammar of the English Language*, Longman Group, London.

Department of English
 Seoul National University
 Shillim-dong, Kwanak-ku
 Seoul 151-742
 Korea