

Osmotic Pressure and the Structure of English

Nahm-Sheik Park
(Seoul National University)

Park, Nahm-Sheik. (2002). Osmotic pressure and the structure of English. *Language Research* 38(2), 473-562.

Building upon the main thesis of Park (1992, 2001), this paper deals with the tendency toward compensatory equilibrium between equivalent structures in English from the “holistic” perspective of traditional grammar. The tendency here involves linguistic additions and deletions getting counterbalanced by surface contractions and expansions, respectively. The phenomenon discussed here is not unlike osmosis in the physical world in that it involves a patterned variation of linguistic elements relative to what they are in construction with. This phenomenon manifests itself on various planes of the structure of English, i.e. in the phonology, orthography, (morpho)syntax and semantics of the language. The main thrust of our discussion here is apparently applicable not just to English but also to Korean and other natural languages so that it may very well qualify as a linguistic universal.

Key words: osmotic, osmosis, compensation, structure, equilibrium

1. Introduction

The mass/space of structural elements in English seems to vary inversely with that of what they are in construction with, either immediately or cumulatively. Underlying this variation appears to be a quest for parity or equilibrium between related or cognate elements of the comparable structural category. We may use the following pair of paraphrase sentences to illustrate our point here.

- (1) a. I've got *some money*.
- b. I've got *some*.

Since (1a) gives rise to (1b) and is thus equivalent thereto, the direct object *some* in the latter arguably absorbs the mass/space for the longer direct object *some money* in the former. Thus *some* in (1b) gets the

mass/space not just for *some* but also for (the trace of) *money* in (1a). *Some* in (1b) is a ‘fattened’ or ‘expanded’ version of *some* in (1a), so to speak, *some* in (1b) being roughly on a par with *some money* in (1a) semantically and thus “phonologically” as well. Thus the mass/space for *some* is evidently in inverse proportion to what it is in construction with in the direct-object noun phrases in question here. The expansion of mass/space via absorption under observation here is apparently attributable to something like osmotic pressure.

Note that, of the two words in question here, the determiner *some* may occur with a great diversity of nouns, whereas the noun *money* may occur with a highly restricted number of determiners only. Thus *some* makes far more cumulative osmotic concession of meaning and pronunciation to *money* and other nouns than *money* does to *some* and other determiners. This is arguably why in a sentence like (1a) the determiner *some* is normally outweighed by the noun *money* both semantically and phonologically.

This paper focuses on osmotic pressure of the sort illustrated by examples like (1) above, as it applies to the structure of the English language. As such, the main theme throughout the paper is osmotic pressure as it throws light on various aspects of the language with respect to its phonology, orthography, (morpho)syntax and semantics.

2. Phonological Osmosis

Elements of English phonology, indeed, appear to be subject to osmotic pressure from what they are in construction with. Our discussion here will focus on the structure of the syllable with specific reference to nucleus, coda and onset. Let us begin our discussion by citing the following minimal pairs.

- (2) a. bid : bit
 b. cab : cap
 c. dog : dock

Note that one and the same nucleus phoneme is phonetically heavier/longer in the first member of each pair here than in the second. Note further that the coda is phonetically more substantive in the second

member of each pair than in the first. This is because, other things being equal, a voiceless consonant outweighs its voiced counterpart in terms of phonetic mass/space.

If this is correct, then we may argue that one and the same nucleus gets phonetically heavier before a lighter coda than before a heavier one. What we have just observed may be summed up as: "The heavier the nucleus, the lighter the coda, and vice versa." Thus in each word pair above, the phonological mass/space of the nucleus varies in inverse proportion to that of the coda with each osmotically counterbalancing the other, so to speak.

Needless to say, an identical account is applicable to why a phonemically identical nucleus is phonetically longer/heavier in the first member of each cognate pair below than in the second member of the same pair.

- (3) a. save : safe
 b. house [hauz] : house [haus]
 c. wreathe : wreath

We also can explain along similar lines why one and the same nucleus phoneme is phonetically a bit heavier/longer in the first member of each word pair below than in the second member of the same pair.

- (4) a. kiss : kit
 b. has : had
 c. leaf : leap
 d. loaves : lobes
 e. heath : heat
 f. breathe : breed

Since a fricative is phonetically lighter than its plosive counterpart, the coda is lighter in the first member of each word pair above than in the second. Note that the nucleus phoneme here is phonetically heavier/longer before the lighter (fricative) coda than before the heavier (plosive coda), which is in tune with our current discussion on phonological osmosis in the English syllable.

Interpretable along essentially identical lines is the patterned allophonic variation of the nucleus observable in each word pair below.

- (5) a. crush : crutch
 b. wish : witch
 c. mash : match

Since a fricative is phonetically lighter/shorter than its affricate counterpart, the coda is lighter/shorter in the first member of each word pair here than in the second. Note that the nucleus is phonetically heavier/longer in the first member of each pair above than in the second. Thus we can say that the heavier/longer nucleus is followed by the lighter/shorter (fricative) coda in the first member of each pair, while the lighter/shorter nucleus is followed by the longer/heavier (affricate) coda in the second member of the same pair. This alternation is again in compliance with the main thrust of our current discussion.

Similarly analyzable is the allophonic variation of the nucleus illustrated by each word pair below.

- (6) a. pen : pend
 b. pass : past
 c. hole : hold

In each word pair here, the nucleus is phonetically heavier/longer before the monoconsonantal coda (i.e. in the first member) than before the biconsonantal coda (i.e. in the second member). Since a monoconsonantal coda is lighter/shorter than its biconsonantal counterpart, the alternation observed here is again corroborative of our claim about osmotic pressure on the structure of the English syllable.

Just as the nucleus varies osmotically as a function of the coda, so does the coda as a function of the nucleus, as can be seen from the difference in the phonetic mass/space of the coda between the two members of each pair below.

- (7) a. fit : feet
 b. hop : hope
 c. look : Luke
 d. tell : tale
 e. fin : fine
 f. moss : mouse

Note that the nucleus is “phonemically” lighter/shorter in the first member of each word pair here than in the second, whereas the coda is phonetically heavier/longer in the first member than in the second. Thus one and the same coda is phonetically heavier/longer after the lighter/shorter nucleus in the first member of each pair here than after the longer/heavier nucleus in the second member of the same pair. This is clearly another piece of evidence for osmotic pressure at work in the English syllable.

The following two sets of words afford us an interesting historical example of osmotic interaction between nucleus and coda in the stem-final syllable.

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

We know from the history of the English language that the word-final alveolar stops /t/ and /d/ here, which had not been there originally, were added later on, probably to round out the syllable in question. Note here that the voiceless /t/ is heavier than the voiced /d/ and that the heavier /t/ was added following a monophthongal stem-final nucleus and the lighter /d/ following a diphthongal stem-final nucleus. Since a monophthong is lighter/shorter than a diphthong, this is in line with our current discussion on phonological osmosis.

Our discussion here also casts light on another example of historical relevance provided by the two word sets below.

- (9) a. rough / cough / tough
 b. dough / bough / plough

The word-final *-gh* here, originally the fricative /x/, has diverged into /f/ in the first set of words above and into zero in the second set, so that the originally identical coda has become heavier in the former than in the latter. Note that the nucleus here is monophthongal when the coda is heavier, i.e. /f/, while it is diphthongal when the coda is lighter, i.e. zero. Thus we can see that the lighter (monophthongal) nucleus is followed by the heavier coda, i.e. /f/, while the heavier (diphthongal) nucleus is followed by the lighter coda, i.e. zero. This again is in line with our current discussion on phonological osmosis in English.

Osmotic pressure also works inside a coda, as can be seen from an examination of word pairs such as (6), repeated below as (6').

- (6') a. pen : pend
 b. pass : past
 c. hole : hold

The coda is monoconsonantal in the first member of each pair here and biconsonantal in the second, with the two codas sharing one and the same consonant phoneme, i.e. /n/ in (6'a), /s/ in (6'b) and /l/ in (6'c). Note that the shared consonant here is normally longer/heavier phonetically in the first member of each pair than in the second. For it may take up all the phonological mass/space for coda in the first member, as opposed to only part of the same mass/space in the second member.

Note incidentally that the same argument applies to one and the same consonant being phonetically longer in a coda consonant-cluster in the environment of a voiced consonant than in that of its voiceless counterpart. We can thus provide a rational account for /n/ and /l/ being phonetically longer in *send* and *build* than in *sent* and *built* respectively, for example.

Osmotic pressure is similarly operative inside an onset, as can be seen from an examination of word pairs such as the following.

- (10) a. pin : spin
 b. till : still
 c. kin : skin

The onset is monoconsonantal in the first member of each pair here and biconsonantal in the second. Note that the two onsets here share a plosive consonant, which is more fortis/ aspirated and thus heavier/ longer in the first member of the pair than in the second. This is because the shared plosive in question monopolizes the phonological mass/space for onset in the first member of each pair, while it has to share the same mass/space with another consonant in the second member of the same pair, i.e. with the onset-initial /s/. This certainly provides an interesting vantage point from which to discuss the dichotomy between fortis/ aspirated and lenis/unaspirated allophones in English.

It is interesting that our current discussion also provides a principled explanation for the allophonic variation of the syllable-initial /s/ in the following pairs of words.

- (11) a. sin : spin
 b. sill : still
 c. sin : skin

Note that the fricative /s/ has to share the onset space with a plosive consonant in the second member of each pair above, but not in the first, which arguably is why it is more fortis and is thus heavier/longer in the first member than in the second.

Needless to say, an essentially identical explanation applies to the syllable-initial alveolar plosive being phonetically softer (or more lenis) in the first member of either pair below than in the second member of the same pair.

- (12) a. tried : tied
 b. dried : died

As we can see from the examples we have just considered in the immediately preceding paragraphs, a consonant is heavier/longer when it occurs by itself than when it occurs as part of a consonant cluster, be it in an onset or in a coda. Needless to say, the phonological mass/space of a consonant in an onset/coda cluster is normally in inverse ratio to that of the rest of the same onset/coda. Thus of the three tokens of /p/ in *pay*, *pray* and *spray*, for example, the first would normally be heavier/longer than the second, which in turn would normally be heavier/longer than the third.

Numerous examples of osmotic pressure at work in English phonology are provided by so-called vocalic laxing, as illustrated by word pairs such as the following.

- (13) a. bone : bonfire
 b. crane : cranberry
 c. produce : product
 d. heal : health
 e. five : fifth
 f. profound : profundity

In each pair here, the nucleus of the (shared) stem-final syllable is (tenser and thus) heavier/longer in the first member than in the second. Note that this (nucleus of the shared) stem-final syllable is followed by extra material (normally in the form of a suffix) in the second member of each pair, but not in the first member. Thus we may argue that the originally tense and heavy nucleus in the first member becomes lax and light in the second member because it has to make osmotic concession of phonological mass/space to the extra material that follows.

Our discussion here evidently affords us a better insight into the stem-nucleus alternation illustrated by the following singular/plural pair.

(14) *breech* : *breeches*

The stem nucleus of the noun here is normally tense in the singular, but always lax in the plural. The obligatory laxing of the stem nucleus observable in *breeches* here apparently results from the osmotic concession of phonological mass/space that the stem makes to the plural suffix. It is interesting to note that the nucleus in *breech* here also is often laxed when the word is used as a verb. This may, in our terms, be due to the osmotic pressure cumulatively exerted on the stem nucleus of the word by the verbal suffixes *-es*, *-ed* and *-ing* as well as the agent-nominal suffix *-er* that often get attached to it.

Incidentally, it may be in order here to point out that the noun *breeches* is a fossilized form so that the plural suffix *-es* is virtually inseparable from the stem *breech*. Thus the plural suffix *-es* is far more tightly bound to the stem in the plural noun *breeches* than, say, in the plural noun *breaches*. As a result, the suffix here exerts far more osmotic pressure on the stem nucleus to get reduced in *breeches* than in *breaches*, which may arguably be part of the reason the nucleus gets laxed in the former, but not in the latter.

Our discussion here offers an interesting angle from which to view the historical development of the stem nucleus vowel in the following words from /ai/ to /i/.

(15) a. *wind* : *windy*

Note that the stem *wind* here combines not just with the suffix *-y* but with a great diversity of other “coda” material, as in *windmill*, *wind-*

break, windblown, window, windstorm, windswept, windward, windless, windshield and *windscreen*. Thus *wind* has been under immense cumulative osmotic pressure to take off its phonological weight one way or another, which it did by getting its original nucleus /ai/ downgraded to /i/. It is interesting that *Jespersen (Part I: 118)* discusses this phenomenon, implying something like osmotic pressure although not in so many words. Incidentally, the nucleus here may sometimes be rendered as /ai/ for the noun *wind*, especially in poetry. This rare retention of /ai/ is not allowed elsewhere, however, because of the greater osmotic pressure exerted by the physically present extra “coda” material.

Given our discussion of (14) and (15), it appears that osmotic pressure is keyed not just to what the affected element is in immediate (physical) construction with but also to what it is in potential (cumulative) construction with. Although taken for granted throughout the paper, this point will be made especially clear in connection with our discussion of semantic osmosis toward the end of Section 4.

It is worth noting at this point that also apparently attributable to osmotic pressure is the variation in the stem-final nucleus in the following set of cognate words.

(16) portray : portrait : portraiture

Note that the stem-final nucleus here is a diphthong in the first word, which is subject to downgrading to a weak monophthong optionally in the second word and obligatorily in the third word. The stem coda, i.e. what follows the stem, being heavier in *portraiture* than in *portrait*, greater osmotic pressure is arguably brought to bear on the stem-final nucleus in the former than in the latter. This is in our terms why the downgrading in question is only optional in *portrait*, while it is obligatory in *portraiture*.

We can explain along similar lines why the stem nucleus is longest in the first member of each cognate-word set below and shortest in the third member of the same set.

(17) a. tide : tidy : tidily
 b. stealth : stealthy : stealthiness
 c. man : manly : manliness

Our discussion here apparently throws light on the variation in the nucleus of the word-initial syllable for the cognates *fiber*, *fibrin* and *fibrillation*. The nucleus here is always /ai/ for *fiber* and either /ai/ or /i/ for *fibril* and *fibrillation* with the proviso that /ai/ is probably more common than /i/ for *fibril* and the other way around for *fibrillation*. This range of variation, which seems to be valid in much of American English, apparently jibes with our current discussion on phonological osmosis in English. For the phonological mass/space of the nucleus in question here varies inversely with that of the “coda” element, i.e. the stretch of phonological material that follows it.

The nucleus of a syllable is osmotically affected not just by what follows but sometimes by what precedes also, as can be seen from examples such as the following.

(18) *migrant* : *immigrant* / *emigrant*

Note that the nucleus of the stem-initial syllable is the strong diphthong /ai/ in *migrant*, which is downgraded to the weak monophthong /i/ in *immigrant* and *emigrant* under osmotic pressure from their prefixal syllables that precede the stem, i.e. from *im-* and *e-* respectively.

Needless to say, similar in nature is the strong-weak vocalic alternation illustrated by the syllable nucleus spelled *o* in the first (positive-negative) cognate pair below and *a* in the second and third pairs.

(19) a. *obey* : *disobey*
 b. *advertent* : *inadvertent*.
 c. *candescent* : *incandescent*

Our discussion here also helps account for an interesting aspect of the nucleic variation illustrated by such word pairs as the following.

(20) a. *room* : *bedroom*
 b. *roof* : *sunroof*

The nucleus for either *room* or *roof* may be either tense or lax with the proviso that the lax nucleus is more likely when the word is part of a compound, as in the second member of each pair here, than when it is not, as in the first member of the same pair. This is apparently because

the nucleus in question has more phonological concession to make when the word in question is in (intra-lexical) construction with another word than when it is not. *Jespersen (Part I: 325)* refers to a similar phenomenon although the notion of osmotic pressure per se is apparently not invoked.

Osmotic pressure also is apparently responsible for the nucleic variations observable in pairs of cognate words such as those given below.

- (21) a. *promo* : promotion
 b. *combo* : combination

The derivation of the first member of each pair here from the second involves upgrading of a vowel from monophthong to diphthong. The vowel affected here is the first-syllable nucleus in the first pair and the second-syllable nucleus in the second pair. We may argue that this nucleus upgrading is designed to osmotically counterbalance the portion deleted in the derivation of the first member of each pair from the second member of the same pair here.

Sometimes more than one nucleus gets upgraded, as can be seen from the following pair of cognate words.

- (22) a. *kilo* : kilogram
 b. *provo* : provisional

Admittedly, the upgrading for the nucleus of the first syllable appears to be optional in the derivation of *kilo* from *kilogram* for a small minority of speakers, which would slightly detract from the value of (22a) as an example supportive of our claim here.

Consonantal segments also appear to get similarly upgraded, as can be seen from cognate pairs such as the following.

- (23) a. *bike* : bicycle
 b. *doll* : Dorothy

Note that /s/ and /r/ get promoted to /k/ and /l/ in (23a) and (23b) respectively to osmotically offset the portions deleted in the derivations involved here.

Sometimes more than one consonant may get promoted, as can be seen

from the following pair of cognate words.

(24) Dick : Richard

The derivation of *Dick* from *Richard* involves double upgrading, i.e. the upgrading of the first and third segments to /d/ and /k/ respectively. *Dick* has two other allo-forms in *Rick* and *Rich*, the first of which involves single upgrading, i.e. the upgrading of the third segment to /k/. Although *Rich* does not involve any superficially visible orthographic upgrading, we can argue that it does involve a more subtle kind of upgrading in that, other things being equal, the three segments making it up are phonetically longer (or more prominent) than the first three segments of *Richard*.

It is worth noting here that osmotic pressure may be invoked to explain the nucleus-coda combination patterns for the two sets of words below.

(25) a. depth / width / health
 b. height / drought / sight

There was a time when the nucleus was tense/heavy for all the words here, when the stems in (25b) terminated with the weak (glottal) fricative /x/, and when the word-final suffix was *-th* for the words in (25b) as well. Thus from the very beginning there must have been a substantial coda disparity between words of the (25a) type and those of the (25b) type. For the coda cluster must have been much heavier in (25a) than in (25b) in that the stem-final consonant was much heavier in the former than in the latter. Consequently, the tense/heavy nucleus here must have been under much greater osmotic pressure to get reduced in words of the (25a) type than in those of (25b) type, which is apparently why it got downgraded to lax vowels in the former, but not in the latter.

The stem-final glottal fricative /x/ in (25b) above seems to have gotten deleted early on, resulting in the mute digraph *-gh-* in present-day English. This *-gh-* deletion via silencing was arguably triggered in large part by the constraint on proximate repetition (of kindred sounds) operating on the two proximate fricatives in word-final position here. Incidentally, this constraint is discussed in detail in Park (1983). The suffixal fricative *-th* following this deleted stem-final fricative arguably

got promoted to the plosive *-t* to osmotically compensate for the deletion.

Our current discussion also seems to be of some relevance to the phenomenon of voicing illustrated by the coda clusters in (26) below, as opposed to those in (27).

- (26) a. life + *-s* => lives
 b. leaf + *-s* => leaves

- (27) a. if + *-s* => ifs
 b. cuff + *-s* => cuffs

Note that the addition of the plural suffix results in biconsonantal (and thus heavy) codas for all the nouns here and that the stem nucleus is diphthongal and thus heavy in (26), as opposed to it being monophthongal (and thus light) in (27). If left alone, the heavy coda cluster would end up combining with a heavy nucleus in (26) here, which would render the resulting nucleus-coda combination rather “obese” This being the case, the coda clusters in (26) get voiced and thus lightened, arguably under osmotic pressure. This voicing, to which the coda clusters in (27) are immune, is apparently conducive to keeping the nucleus-coda combination from getting too obese in (26), simultaneously fostering syllable-structure parity between the words of the (26) type and those of the (27) type.

Sometimes the voicing of the coda cluster here does not seem to materialize unless the nucleus in question is patently diphthongal. Thus when the nucleus is a weak diphthong such as /i:/ or /u:/, the voicing of the coda cluster often does not materialize, as can be seen from examples such as the following.

- (28) a. roof + *-s* => roofs
 b. chief + *-s* => chiefs

It may be that the coda cluster does not get voiced here partly because the nucleus-coda combination is not that obese so that the osmotic pressure for it to get voiced may not be that strong.

Needless to say, the coda-cluster voicing operative in (29) below, but not in (30), is essentially identical to that operative in (26) above, but not in (27).

- (29) a. mouth + -s => mouths
 b. truth + -s => truths
- (30) a. death + -s => deaths
 b. cloth + -s => cloths

Note that the coda cluster *-ths* here gets voiced obligatorily only when the nucleus is strongly diphthongal and thus distinctly heavy. Thus the coda cluster voicing is obligatory in *mouths*, but optional in *truths* (as well as in such plural forms as *baths*, *oaths* and *sheaths*).

It may be in order to observe here that the genitive suffix *'s* is less seamlessly attached to the noun stem than is the plural suffix *-s* so that the former does not exert as much osmotic pressure on the stem coda as does the latter. This is evidently why superficially similar coda clusters do not get voiced in genitive forms such as *life's* and *mouth's*. Incidentally, the apostrophe here is phonologically real in that it evidently indicates that there is a gap between the two immediate constituents of a genitive form, which evidently is what blocks the coda cluster in question from getting osmotically voiced.

Incidentally, we may rationalize the /f/-/v/ alternation illustrated by the following cognate pair by resorting to osmotic pressure.

- (31) plaintiff : plaintive

Both words here originated in one and the same Middle English adjective *plaintiff* and the word-final /f/ has since been voiced to /v/ for the adjective, but not for the noun. Note here that an adjective may normally take up only (a peripheral) part of noun-phrase mass/space, whereas a noun can take up most (or all) thereof. Thus the word in question here must have had more phonological space assigned to it as a noun than as an adjective. In light of this, the voicing of /f/ to /v/ here may be viewed as arguably having served to cut the word down to size so as to fit it comfortably into the relatively confined slot of a (premodifying) adjective. As a noun, on the other hand, the word has apparently not faced the same compelling osmotic pressure to get cut down to size via the same /f/-to-/v/ voicing. For the nominal slot is massive/spacious enough to comfortably accommodate the voiceless word-final /f/ here as it is.

Explainable along similar lines is the pronunciation difference for the nucleus of the suffix *-ful* between words of the (32a) type and those of the (32b) type below.

- (32) a. *handful* / *mouthful* / *armful*
 b. *careful* / *youthful* / *beautiful*

As noun-phrase constituents, *-ful* nouns such as those in (32a) normally get more mass/ space assigned to them than do *-ful* adjectives such as those in (32b). For it is generally the case that nouns figure as central constituents of noun phrases and adjectives as mere peripheral constituents of the same. Thus *-ful* gets more phonological mass/space assigned to it in nouns such as those in (32a) than in adjectives such as those in (32b), which is evidently why its nucleus is more prominent in the former than in the latter. Incidentally, we are assuming here that the high back vowel /u/ is heavier than the schwa.

Osmotic pressure also seems to underlie the change from the parenthesized Middle English forms to their respective contemporary counterparts given below.

- (33) a. riddle (<=redel)
 b. kettle (<=ketel)
 c. shuttle (<=shotil)
- (34) a. button (<=boton)
 b. cotton (<=coton)
 c. kitten (<=kitoun)

Note that all the words listed here are disyllabic both in Middle English and present-day English and that the nucleus in the second syllable of each Middle English form here seems to have gotten deleted sometime in the early Modern English period. Note further that the word-final sonorant consonant subsequently became syllabic in apparent compensation for the disappearance of this original second-syllable nucleus, arguably under osmotic pressure.

Osmotic pressure also is evidently instrumental in explaining the two alternative pronunciations of the following word.

(35) Kabul

Either vowel here may be either stressed/tense or unstressed/lax, normally with the proviso that the two may not be both stressed/tense or both unstressed/lax. Arguably triggered by osmotic pressure, this alternation seems to put the two alternative pronunciations on a par in vocalic mass/weight.

We are now in a position to offer an enlightened account for the alternation between the two allomorphs of the English indefinite article, as illustrated below.

- (36) a. an ode : a node
 b. an aim : a name

Note that the heavier/longer allomorph, i.e. *an*, co-occurs with the lighter/shorter (vowel-initial) syllable, i.e. *ode* or *aim*, while its lighter/shorter counterpart, i.e. *a*, co-occurs with the heavier/ longer (consonant-initial) syllable, i.e. *node* or *name*. Thus the historical derivation of *a* from *an* seems to have been motivated at least in part by a need to help establish osmosis-triggered equilibrium for the phonological sequences concerned (as well as to help keep consonants or vowels from colliding head on at syllable/lexeme boundaries).

Needless to say, exactly the same rationale is applicable to the alternations exemplified by such pairs as the following from early Modern English.

- (37) a. mine ode : my node
 b. thine ode : thy node

Osmotic pressure also apparently underlies such prosodic variation as is observable between the prepositions and their particle counterparts that figure in pairs of paraphrase sentences such as those below.

- (38) a. He's *in his office*.
 b. He's *in*.

- (39) a. He came *down the hill*.
 b. He came *down*.

- (40) a. *Keep off the grass.*
 b. *Keep off.*

The second member of each pair here derives from the first member of the same pair with the prepositional phrase in the first member arguably getting kind of compacted into the particle in the second member. Thus the de-prepositional particle in the second member osmotically takes up the phonological mass/space of the entire prepositional phrase in the first member, that is, not just the preposition but also its complement deleted in the derivation involved here. This is apparently why the particle in the second member of each pair here is (semantically heavier and thus) prosodically more prominent than the preposition in the first member of the same pair. Incidentally, a basically identical line of reasoning may be used in explaining why stranded prepositions are generally more prominent prosodically than their non-stranded counterparts.

We can explain along essentially identical lines why “auxiliaries” are normally rendered with greater prosodic prominence clause-finally than elsewhere, as can be seen from an examination of the following pairs of paraphrase sentences.

- (41) a. He *will come here.*
 b. He *will.*

- (42) a. He *has come here.*
 b. He *has.*

- (43) a. He *is American.*
 b. He *is.*

In the derivation of the second member of each pair here from the first, the “auxiliary”-initial, full verb phrase may be assumed to get compacted into the auxiliary-only, contracted verb phrase. Thus the lone auxiliary in the second member gets to osmotically bear the prosodic mass/space of not just the auxiliary in the first member but also the rest of the verb phrase that gets deleted in the derivation involved here. This is arguably why one and the same auxiliary is (semantically and thus) prosodically more prominent when clause-final than elsewhere.

We will have more to say about examples like (38)-(43) in Section 5,

which deals with semantic osmosis in English. For prosodic variation of the sort under discussion here is apparently conditional upon corresponding semantic variation. Incidentally, the “auxiliary” here never gets contracted to forms like *'ll* or *'s* clause-finally apparently because of the extra semantic/phonological mass/space that it has come to bear under osmotic pressure.

3. Orthographic Osmosis

Osmotic pressure seems to be instrumental in resolving quite a few apparent anomalies in English orthography. Among these apparent anomalies is the frequent use of uppercase initials for “relationship” terms when they are not qualified by modifiers such as determiners, as can be seen from the following paraphrase pairs.

- (44) a. I'm waiting for *Dad*.
 b. I'm waiting for *my dad*.

- (45) a. I took *Mom/Mother* to the hospital.
 b. I took *the mom/mother* to the hospital.

- (46) a. This shopping cart has plenty of leg room for *Junior* to ride.
 b. This shopping cart has plenty of leg room for *the junior* to ride.

Note that the relationship term in question has to share the orthographic space for noun phrase with a determiner in the second member of each pair here, but not in the first. The first letter of this relationship term is arguably capitalized in the first member of each pair here in indication of its osmotic absorption of the orthographic mass/space for an entire noun phrase (as its sole occupant). In the second member of the same pair, however, the same term has to make some osmotic concession to another occupant of the noun phrase slot in question, i.e. a determiner, so that there is hardly any room here for an uppercase initial nor is there, for that matter, any need therefor.

Needless to say, explainable along similar lines is the alternation between *Coach* and *coach*, as illustrated by the following paraphrase pair.

- (47) a. I know *Coach* does not approve of things like this.
 b. I know *my coach* does not approve of things like this.

Parenthetically, our discussion here apparently sheds light on the alternation between the uppercase *G* in *God* (“the Creator”) and the lowercase *g* in *god* (“a supernatural being”). Let us take a look at the following pair of sentences illustrating the complementary distribution between *God* and *god*.

- (48) a. I am afraid of scientists playing *God*.
 b. Mars was *the ancient Roman god* of war.

As can be seen here, we may use a full modifier string, not just an article, with *god*, but not normally with *God*. Thus *God* in (48a) takes up the whole slot for the noun-phrase of which it is the head and sole constituent, while *god* in (48b) has to share with a modifier string (i.e. *the ancient Roman - of war*) the slot for the noun phrase of which it is the head but not the sole constituent. We can say that the uppercase initial *G* is resorted to in *God* here, but not in *god*, because the former has larger orthographic mass/space to fill up than does the latter, so to speak.

Admittedly, the *G-g* alternation here may also have something to do with semantics in that “God” refers to a far higher-order divine being than does “a god,” which perspective is taken in Park (1993). We may speculate here that *God* is assigned much larger semantic mass/space than is *god* because the former normally has a much larger semantic slot to fit into than does the latter. For the former, but not the latter, has the semantic slot of a whole noun phrase to occupy.

We may observe in passing that similarly rationalizable is the uppcasing of initials in such other quasi-proper nouns as *the Shepherd* and *the Master* for “Jesus Christ,” *the Virgin* for “the Holy Mother” and *the Book* for “the Holy Bible.” That is, except that these quasi-proper nouns allow predetermination by the definite article. Incidentally, *the Bard* for “Shakespeare” may also be written as *the bard*, which may (or may not) be due to the less than divine status of William Shakespeare. It goes without saying that we can similarly account for the uppcase initial *H* in *the Holocaust*.

We may contend at this point that the rationalization for the

uppercasing of the initial in *God* may be applicable mutatis mutandis to the convention of using uppercase initials for proper nouns in English. Like *God*, a proper noun does not normally have to share with a qualifying modifier the orthographic mass/space of the noun phrase of which it is the head. Thus it also has plenty of orthographic mass/space to osmotically absorb, which it apparently does by getting its initial capitalized.

Our discussion here also apparently helps explain why the names of the days of the week (e.g. *Sunday* and *Monday*) or of the months of the year (e.g. *January* and *February*) begin with uppercase letters. We may argue that these words are quasi-proper nouns and thus normally used without any qualifying modifiers so that they have plenty of orthographic mass/space to digest, so to speak, which they do by resorting to uppercase initials.

We can apparently explain along identical lines why quasi-proper nouns of recent coinage such as those italicized in the sentences below begin with uppercase letters.

- (49) a. Most of my friends are on *the Internet*.
 b. Most of my friends are on *the (')Net*.
 c. Plenty of information is available on *the Web*.

It is interesting to note that an initial thus upgraded may sometimes get downgraded back to a lowercase letter when something extra is added to the quasi-proper noun in question, as can be seen from the following paraphrase pair.

- (50) a. I got some valuable information from their *Web site*.
 b. I got some valuable information from their *web site*.

The optional *W-to-w* downgrading phenomenon illustrated above is arguably designed to mark the osmotic concession of orthographic mass/space made to the addition of *site* to the quasi-proper noun *Web*.

Note that the orthographic downgrading under discussion here is obligatory when the (compound) word in question is written *website*, that is, with no space between its two immediate constituents. This may be because the word is perceived more distinctly as a single word when there is no “visible” space between its two constituents. Thus the first

constituent here may be under greater osmotic pressure to make an orthographic concession to the second in *website* than in *Web site*, which is apparently why the downgrading of the initial is obligatory in the former, but not in the latter. The word *Web page*, on the other hand, is always written with space between its two immediate constituents (at least for now), so that the word-initial letter may only optionally get downgraded to the lowercase *w*, resulting in *web page*.

The line of reasoning being followed here apparently makes for a principled discussion of why the italicized noun may begin with either *E* or *e* in the first paraphrase pair below, while it may begin only with *e* in the second paraphrase pair.

- (51) a. *The Earth* is round.
 b. *The earth* is round.
- (52) a. What *on earth* is going on here?
 b. *What *on Earth* is going on here?

The noun in question here is an independent word functioning as a quasi-proper noun in the first pair, but not in the second pair. It is not an independent word in the second pair because it is part of an idiom, i.e. *on earth*, which is functionally monolexemic. In fact, the intensifier *on earth* here may be regarded as an (idiomatic) compound word. This being the case, the noun in question apparently has less concession of orthographic mass/space to make in the first pair than in the second. This may be part of the reason why the word may begin with either *E* or *e* in the first pair, but only with *e* in the second pair. Note here that we can now easily explain why *earth* always begins with the lowercase *e* when it forms part of an idiom such as *down to earth*, *the salt of the earth*, *to move heaven and earth*, or the like.

Note in this connection that the idiomatic compound word *on earth* in (52) is different from the non-idiomatic phrasal combination *on earth/Earth* illustrated in the following paraphrase pair.

- (53) a. Okinawans live longer than anyone else *on earth*.
 b. Okinawans live longer than anyone else *on Earth*.

The word *earth/Earth* here is a quasi-proper noun referring to our home

planet, just as it is in (51) above, which is why the initial may be either *e* or *E*. Note in this connection that *earth* is in much closer construction with *on* in (52) than in (53), which is why *earth* ends up making much more osmotic concession of orthographic mass/space to *on* in the former than in the latter. As we shall see in Section 5, *on* and *earth* here also make more mutual concession of meaning when *on earth* is an idiomatic compound word than when it is a non-idiomatic phrasal combination.

Incidentally, *hell* and *heaven* also are normally quasi-proper nouns that may begin with either *h* or *H*. They must begin with the lowercase *h* only, however, as part of idiomatic compound words, as can be seen from the following paraphrase pairs, in which both *the hell* and *for heaven's sake* here are idiomatic compounds used as intensifiers.

- (54) a. Where *the hell* have you been?
 b. *Where *the Hell* have you been?
- (55) a. Where have you been, *for heaven's sake*?
 b. *Where have you been, *for Heaven's sake*?

Thus we can see that both *hell/Hell* and *heaven/Heaven* display the same traits of orthographic behavior as *earth/Earth* with respect to their initials.

Our discussion of the uppercasing phenomenon here may shed some light on the interesting alternation between *John* and its cognate *john*, as illustrated in the following pair of sentences.

- (56) a. She was talking to *John*.
 b. She was talking to *a john*.

We know from the history of English that the common noun *john* was derived via conversion from the proper noun *John* sometime during the 1910s. What is of interest to us is that the word-initial letter here is downgraded from *J* to *j* apparently to make osmotic room for a qualifier/determiner, with which the converted (common) noun normally occurs.

It is interesting to note that *john* in the sense of “a prostitute’s customer” may, albeit rather rarely, be spelled as *John* with an uppcase initial, while the same word in the sense of “a men’s room” is always

spelled *john* with a lowercase initial. It may be the case that the common-noun usage of the word is somehow less well established in the former sense than in the latter. Or, the human (thus more literal) reference of *John/john*, as opposed to its non-human (and thus less literal) reference, may have something to do with the relative degree to which the orthographic downgrading in question here is allowed. We shall, in fact, suggest this latter perspective in connection with *Herculean/herculean* and *Protean/protean* later on.

An essentially identical interpretation is applicable to the alternation between uppercase and lowercase initials for the italicized nouns in each sentence pair below.

- (57) a. Bob felt attracted to *Jane*.
 b. Bob felt attracted to *a new jane*.
- (58) a. John is dating *Jill*.
 b. John is dating *another jill*.
- (59) a. He was cast as *Lothario* in the play.
 b. He was cast as *a lothario* in the play.

Incidentally, *jill* and *lothario* may be replaced by *Jill* and *Lothario* respectively for some users of English, which may indicate that they may still feel quite strongly about the proper-noun origins for the common nouns *jill* and *lothario*, whose “common-nounization” has not yet been completed.

The lowercasing of the sort under discussion here is also observable when proper nouns are converted to other parts of speech such as the verb and adjective. Let us take the following, for example.

- (60) The hungry puppy has hoovered up all the kibble in the bowl.

The verb *hoover(ed)* here comes from the trademark of a vacuum cleaner manufacturer, i.e. *Hoover*, (and sometimes retains its uppercase initial). Note that this verb is in construction with the direct object *all the kibble in the bowl* here (and in potential construction with a great diversity of other object noun phrases elsewhere), to which it has to make some osmotic concession of orthographic mass/space. This is, in our terms, why

the verb here gets its initial demoted from *H* to *h*.

Another example of *H-to-h* demotion is afforded by the following sentence, in which *hoover* is an adjectival (noun) derived from the proper noun (*Herbert Clark*) *Hoover*.

- (61) *The hoover apron* was never really in vogue.

Note that the adjectival *hoover* is in construction with *The - apron* and thus has to osmotically share the orthographic slot for the noun phrase therewith, which is apparently what occasions the *H-to-h* downgrading in question.

Additional examples of lowercasing are afforded by the italicized words in the following sentence pairs.

- (62) a. The crutch of time does more than the club of *Hercules*.
 b. Digging the tunnel was a *herculean* task.
- (63) a. *Proteus* was a sea god capable of changing from one form to another.
 b. Are all women really so *protean*?

It may be pointed out that the orthographic downgrading of the initial here seems to be blocked when the reference to *Hercules* or *Proteus* is literal, rather than figurative. This may have something to do with either adjective here being more “proper-adjectivish,” as it were, when the reference is literal than when it is figurative.

Interestingly enough, our discussion here is of apparent relevance to the “free” variation between *the North* and *the north* in the following paraphrase pair.

- (64) a. Hundreds of thousands of people came down from *the North* for the holidays.
 b. Hundreds of thousands of people came down from *the north* for the holidays.

We may assume that *the North* here derives from *the north*, either form ultimately deriving from *the north of a particular country (or region)*. Based on this assumption, we may argue that the optional *n-to-N*

upgrading here is designed to osmotically compensate for the deletion of *of a particular country (or region) from the north of a particular country (or region)*. It goes without saying that *south* in *the south*, *east* in *the east* and *west* in *the west* may get their initials upgraded to their respective uppercase letters for precisely the same reason.

Note in this connection that the *n-to-N* upgrading is obligatory in *up North*, while it is optional in *up in the north/North* or *up to the north/North*, as can be seen from the following paraphrase pairs.

- (65) a. Brad is from somewhere *up North*.
 b. Brad is from somewhere *up in the north/North*.
- (66) a. Brad is thinking of going *up North*.
 b. Brad is thinking of going *up to the north/North*.

Up North here apparently originates in *up in the north/North* and *up to the north/North*, which in turn apparently originate in *up in the north of a particular country (or region)* and *up to the north of a particular country (or region)*, respectively. We may thus argue that the *n-to-N* upgrading in the derivations involved here is obligatory because some osmotic compensation must be made not just for *of a particular country (or region)* but also for *in the* or *to the*, as the case may be.

Needless to say, interpretable along identical lines is the obligatory uppcasing of the underlined initial in the first member of each paraphrase pair below, as opposed to the optional uppcasing thereof in the second member of the same pair.

- (67) a. We are going to move *down South*.
 b. We are going to move *down to the south/South*.
- (68) a. Glen went to college *back East*.
 b. Glen went to college *back in the east/East*.
- (69) a. They were going *out West*.
 b. They were going *out to the west/West*.

Osmotic pressure also apparently lies at the root of the tendency to capitalize an acronym in its entirety, as is illustrated by paraphrase pairs

such as the following.

- (70) a. ATM <= automated teller machine
 b. GIGO <= Garbage in, garbage out
 c. BYOB <= Bring your own bottle.
 d. IPO <= initial public offering
 e. EMT <= emergency medical technician

Each constituent letter of an acronym arguably is written with an uppercase letter to show that it has osmotically taken up the mass/space of an entire word for which it stands. The first letter in *ATM*, for one, is promoted from *a* to *A* to show that it stands not just for the initial letter *a-* of *automated* but for the rest of *automated*, i.e. for *-utomated*, as well.

Needless to say, explainable along similar lines is the uppercasing in quasi-acronyms such as the following.

- (71) a. ID <= *identification/identify*
 b. IV <= *intravenous*
 c. TB <= *tuberculosis*
 d. TX <= *Texas*
 e. A-bomb <= *atomic bomb*
 f. super-G <= *super giant slalom*

Our discussion here appears to be instrumental in accounting for the apparent anomaly of the first-person singular pronoun being spelled with the uppercase *I* in present-day English, which normally took either of the following forms in Old and Middle English.

- (72) *ik* / *ich*

The historical development of *ik* or *ich* into *I* involves an acronymic abbreviation of sorts so that the use of the capital *I* here is arguably designed to help osmotically make up for the orthographic mass/space lost to the deletion of *-k* from *ik* or of *-ch* from *ich*.

Admittedly, the first-person singular pronoun seems to have had a third form, i.e. *i*, briefly in (late) Middle English. Since *i* was presumably used only rarely and temporarily, probably as a makeshift bridge of sorts

between *ik/ich* and *I*, we will ignore this form in our discussion here as a mere transient form. Besides, even this bridge form *i* was sounded as a tense high-front vowel, which was a phonologically upgraded version of its lax counterpart in the earlier form *ik/ich*. Note that this nucleic upgrading also was apparently triggered by osmotic pressure. Note also that the eventual phonological upgrading of the nucleus of the pronoun here from /i/ in *ik/ich* to /ai/ in *I* may be cited in support of phonological osmosis and can thus just as well belong in the immediately preceding section of the current paper.

Rationalizable along essentially identical lines is the noun-initial uppercasing involved in the derivation of the first member of the following pair from the second member of the same pair.

(73) the Continent <= the European continent

The *c*-to-*C* upgrading here evidently serves to osmotically make up for the deletion of *European*.

Block language represents another interesting case of recourse to uppercase letters as a means of osmotically compensating for what gets deleted in the process of deriving “abbreviated” forms. Book titles and chapter headings, for example, normally have at least their initials capitalized, so as to make up for whole blocks of suppressed language or information that they stand for.

It is interesting to note in this connection that, other things being equal, uppercase letters apparently often figure far more prominently in book titles than in chapter headings. Let us take, for example, a famous book by British physicist Stephen Hawking, which is titled thus.

(74) *A BRIEF HISTORY OF TIME*

Note that the whole title is capitalized here, whereas the chapter headings have only the initials of their major constituent words capitalized, as can be seen from the headings for the first three chapters of the book repeated here as (75) below.

- (75) a. Our Picture of the Universe
 b. Space and Time
 c. Our Expanding Universe

In our terms, book titles are often far more heavily capitalized than chapter headings because the former compensate for far larger blocks of suppressed language/information than do the latter.

Our discussion here shines an interesting light on the uppercase-lowercase alternation illustrated by the two italicized paraphrase sets of nouns in the examples given below.

- (76) a. I want to be a *disc jockey*.
 b. I want to be a *DJ*.
 c. I want to be a *deejay*.
- (77) a. The *master of ceremonies* did a superb job.
 b. The *MC* did a superb job.
 c. The *emcee* did a superb job.

Note that the initials get uppercased here when the underlying compound noun in each paraphrase set here is contracted into an acronym. Note further that this “acronym” in its entirety gets back down to lowercase letters when it is de-acronymized by getting spelled out as pronounced and thus orthographically re-expanded. This range of spelling fluctuation is arguably the result of orthographic osmosis applying to the paraphrase sets of nouns in question here.

Our current discussion also seems to make available a rational account for the alternation between *the 49ers* and *the Niners* in the following paraphrase pair.

- (78) a. The Falcons finally beat *the 49ers* 39 to 19.
 b. The Falcons finally beat *the Niners* 39 to 19.

The San Francisco-based NFL franchise in question here is normally referred to in writing as *the (San Francisco) 49ers*. The name of the team may be further reduced by deleting 4 from 49, the result of which is normally spelled out as *the Niners*, seldom as *the 9ers*. The spelling out of 9 as *Nine* here is arguably intended to compensate for the deletion of 4, thereby making it possible for the two alternant forms involved here achieve and maintain an osmotic equilibrium of the sort under discussion here.

Needless to say, an exactly identical account is applicable to the

orthographic alternation between *the 76ers* and *the Sixers*, as illustrated in the following paraphrase pair.

- (79) a. Iverson was picked by *the 76ers* in the first round of the NBA draft.
 b. Iverson was picked by *the Sixers* in the first round of the NBA draft.

Similarly rationalizable may be the orthographic variation between *1992* and its alternant *ninety-two* in the paraphrase pair below.

- (80) a. They say that *1992* was a sad year for the Bush family.
 b. They say that *ninety-two* was a sad year for the Bush family.

Referred to in full, the year in question is normally written in Arabic numerals in its entirety, i.e. as *1992*. When the expression is reduced by deleting the first two digits, i.e. *19*, on the other hand, it is usually spelled *ninety-two* or *'92*. The spelling out as *ninety-two* here may be motivated by a desire to compensate for the deletion of the first two digits from *1992* and thus help the two alternant orthographic forms achieve and maintain osmosis-conditioned parity. Incidentally, the apostrophe in *'92* also arguably serves to osmotically make up for the deleted two digits here. The use of punctuation marks such as the apostrophe to compensate for deleted elements is discussed in some detail elsewhere (Park 1992).

It is interesting to observe at this point that single-digit numerals, but not multiple-digit numerals, tend to be spelled out (except in sentence-initial position), especially in journalistic writing, as can be seen from sentences such as those cited below.

- (81) a. O'Neal had *32* points, *11* rebounds and *five* blocks.
 b. It could take *five* to *10* years.
 c. She is survived by *five* daughters and *12* grandchildren.

Suppose, for the sake of argument, that both multiple-digit and single-digit numerals here were all either written in Arabic numerals or fully spelled out. Then we would end up with a considerable orthographic disparity between these two kinds of numerals. Thus the

single-digit numerals here may get spelled out in full, while their multiple-digit counterparts do not, so as to arguably help bridge or block a potential orthographic gap of this sort.

It may also be noted here that written English allows only Arabic numerals to occur with such “affixal” forms as \$, %, *km*, etc., as in \$25, 98%, 42*km*, etc. This is apparently because the numerals must make some intralexical osmotic concession to the affixal forms in question here (so as to block formation of overly obese words). If the numerals were spelled out here, the resulting words would be osmotically way out of line, so to speak.

Sometimes osmotic pressure results in the doubling of a letter, as can be seen from the following paraphrase pair, the first member of which is derived from the second.

(82) soccer <= association football

Note here that *association football* is abbreviated to *soccer*, in the course of which the letter *c* gets doubled to *cc* to help make up for the portion deleted from the longer source form, i.e. for *as-iation football*. We might say that the extra letter *c* here (in tandem with *-er*) osmotically accommodates the orthographic mass/space originally occupied by the portion deleted in the derivational process involved. Parenthetically, the upgrading here is not just orthographic but phonological as well. For it so happens that the /s/ for *-c-* in *association football* is phonologically upgraded to the /k/ for *-cc-* in *soccer*.

Construable along essentially identical lines is the phenomenon of letter gemination illustrated by paraphrase pairs such as the following, in each of which the first member is derived from the second.

- (83) a. *divvy* <= *dividend*
 b. *civvies* <= *civilian clothes*
 c. *tummy* <= *stomach*
 d. *jammy* <= *pajama*
 e. *granny* <= *grandmother*
 f. *hubby* <= *husband*
 g. *innards* <= *inwards*
 h. *igg* <= *ignore*
 i. *veg(g)ie* <= *vegetable*
 j. *preemie* <= *premature baby*

Granted, the letter doubling here also often serves the ancillary purpose of keeping intact the original phonological value of the immediately preceding vowel, as in *tummy* (<=*stomach*) and *granny* (<=*grandmother*).

Our current discussion makes available a principled explanation for the derivation of the first member of the following paraphrase pair from the second member of the same pair.

(84) *iff* <= *if and only if*

The second token of *f* in *iff* here is in our terms an osmotically-triggered compensatory trace for *and only if* in *if and only if*.

Osmotic pressure also may be invoked in explaining the phenomenon of orthographic “de-clustering,” as exemplified by the spelling alternation for the stem-nucleus shown in the following set of words.

- (85) a. float
 b. flotation / floatation
 c. flotage / floatage

The (underlying) stem-nucleus digraph *-oa-* in *float* normally gets de-clustered to *-o-* when the noun-forming suffix *-ation* or *-age* is added. This de-clustering is evidently designed to make an osmotic concession of orthographic mass/space to the addition of the suffix in question. Interestingly, the orthographic *oa/o* alternation here is phonetically real in that the longer form *oa* in *float* tends to represent a phonetically longer sound than does the shorter form *o* in *flotation* and *flotage*.

Note in this connection that the de-clustering under discussion here does not occur when the suffix in question is a productive one such as *-ing*, *-ed*, *-s*, *-able*, *-er* or *-ee*, as can be seen from the following set of forms derived from *float*.

- (86) a. floating = *floting
 b. floated = *floted
 c. floats = *flots
 d. floatable = *flotable
 e. floater = *floter
 f. floatee = *flotee

Less productive suffixes (such as *-ation* and *-age*) seem to be more tightly bound to the stem than do more productive ones (such as *-ing* and *-ed*) with the result that the former exert heavier pressure for osmotic concession on the stem to which they are suffixed than do the latter. For a more detailed discussion of the phenomenon under discussion here, see Park (1992).

It is interesting to note in this connection that the osmosis-motivated variation in the relative phonological mass/space of a syllable segment also is often orthographically reflected as can be seen from examples such as the following.

- (87) a. tack : take
 b. lick : like
 c. sock : soak
 d. puck : puke

The nucleus is lighter/shorter “phonemically” in the first member of each word pair here than in the second. On the other hand, the coda is phonetically heavier/ longer in the first member of each pair than in the second. Thus the lighter/shorter nucleus combines with the heavier/ longer coda in the first member of each pair here, and vice versa in the second member of the same pair.

Note here that, of the two consonantal codas in each word pair above, the phonetically heavier/longer one is represented by the digraph *ck* and the phonetically lighter/shorter one by the single letter *k*. Note also that, of the two vocalic nuclei in each word pair here, the heavier/longer one is spelled with two vowel letters and the lighter/shorter one with just one vowel letter. Incidentally, the two vowel letters for the heavier/ longer nucleus here may be either continuous, as in *soak*, or discontinuous, as in *take*, *like* and *poke*. At any rate, what is important here is that this is evidently an orthographic reflex of the phonological osmosis at work here.

Osmotic pressure also apparently bears on the orthographic stem-coda alternation in cognate word pairs such as the following.

- (88) a. pass : past
 b. shall : shalt
 c. spill : spilt

Originally comprising two consonant letters, the stem coda loses one of these two letters when the suffix *-t* is added. The loss of the letter in question here is arguably designed to make osmotic room for the addition of the suffix. Note that the stem-coda alternation under discussion here is an orthographic reflex of the underlying allophonic variation of the stem codas involved, with *ss* and *ll* representing longer underlying allophones than do *s* and *l* respectively.

It is interesting that the stem-final consonant letter here does not get deleted when the suffix *-s* is added, as can be seen from examples such as the following.

- (89) a. spill : spilt : spills
 b. spell : spelt : spells

The suffix *-s* represents a far softer/lighter sound than does the suffix *-t* so that the latter takes up far more phonetic (and thus orthographic) room in the coda here than does the former. Thus the suffix *-t* is massive enough to squeeze out the stem-final consonant letter here, while the suffix *-s* is not massive enough to do so.

Incidentally, we may get *spilled* (but not **spild*) as the regular past (participle) form of *spill* probably because /d/ is a softer/lighter sound than is /t/. It may also be the case that the past (participle) suffix /t/ is more closely bound to the verb stem here than is its variant in /d/ with the result that the former may exert greater osmotic pressure on the stem coda than does the latter.

Examples like (85)-(89) do indeed go to show that English orthography manifests an unusually high degree of phonological reality. We may further drive home the point here by citing additional examples of relevance such as the following.

- (90) a. pool : pull
 b. feel : fill

Note that the nucleus is “phonemically” heavier/longer in the first member of each pair here than in the second, while the coda is phonetically heavier/longer in the second member than in the first. Note further that the heavier/longer segment is spelled with a double letter and the lighter/shorter segment with a single letter. Also, the nucleus

varies inversely with the coda here both phonologically and orthographically, which is in compliance with the main thrust of our claim about osmotic pressure on the phonological/orthographic structure of the English syllable.

Note parenthetically that osmotic pressure is instrumental in explaining why the first member of each pair below is a more natural orthographic word in English than the second member of the same pair.

- (91) a. pool : ??pooll
b. pull : (?)pul

- (92) a. feel : ??feell
b. fill : (?)fil

It may be observed here that a typical monosyllable in English is such that, orthographically, the mass/space of the nucleus normally varies inversely with that of the coda. This basic orthographic canon of the English syllable is complied with by the first member of each word pair above, but not by the second member of the same pair, which is why the former is an acceptable orthographic word, but not the latter.

It is interesting to note in this connection that overweight syllables appear to disturb osmotic equilibrium far more seriously than do underweight syllables. In fact, an overweight monosyllable like ??doong is an orthographic monster, whereas an underweight monosyllable like (?)dil may qualify, albeit marginally, as an acceptably spelled word. It may be that slim is better than fat in English orthography just as it apparently is in the real world.

4. Formal Osmosis

Our present discussion can also be instrumental in accounting for the formal difference in object-clause complementation for *to learn* and *to know*, as shown in the following paraphrase pairs.

- (93) a. He has *learned how to* dance.
b. He has *learned to* dance.

- (94) a. He *knows how to* dance.
 b. *He *knows to* dance. ("He *can* dance.")

Note that the dynamic *to learn* is heavier (in meaning) than the stative *to know* and that either the heavier *how to* or the lighter *to* is allowed as object clause complementizer for the former, as opposed to only the heavier *how to* for the latter. This alternation, which is apparently osmotically triggered, may be designed to help the *learn-* and *know-* initial verb phrases here to better attain rough formal parity.

A similar rationalization is apparently applicable to the alternation between *had* and *was at* in the following paraphrase pair.

- (95) a. Brad Radke *was at* a 2.49 ERA in his last three starts.
 b. Brad Radke *had* a 2.49 ERA in his last three starts.

We assume here that *had* is heavier than *was* so that we need to supplement the latter with the preposition *at* so as to put it on a formal par with the former. Incidentally, the meaning of "having/possession" appears to outweigh that of "being/existence," which is why we assume here that *had* is heavier than *was*. Apparently amenable to a similar analysis is the presence of the preposition *of* in *to be of value/importance/significance*, as opposed to its absence from the cognately equivalent *to have value/importance/significance*.

We can probably explain along basically identical lines, at least in part, why the preposition *in* figures in the verb phrase of the first member of each sentence pair below, but not in the second.

- (96) a. I *was in* charge of the business section of the company.
 b. I *took charge of* the business section of the company.
- (97) a. Friendly forces *were in* control of the fortress.
 b. Friendly forces *gained control of* the fortress.
- (98) a. The patient *was in* pain.
 b. The patient *was feeling* pain.

It is interesting to observe at this point that there was a time when the particles *in* and *out* had their respective variants *inne* and *oute*, with the

variation being apparently osmosis- triggered. The two alternants of either particle here were in complementary distribution with each other, as can be seen from sentence pairs such as the following.

- (99) a. He *was inne*.
b. He *came in*.

- (100) a. He *was oute*.
b. He *went out*.

We can see here that the heavier/longer variant of either particle combines with *to be*, while the lighter/shorter variant of the same particle combines with *to come* or *to go*. Since *to be* is lighter than either *to come* or *to go*, this complementary distribution between the two variants of either particle here is arguably designed to help the verb phrases in the two members of each pair here achieve formal equilibrium.

It is worth noting here that something like the particle alternation illustrated by either (99) or (100) can also be found in present-day English as well, albeit not in exactly the same form. Whereas the two alternants seem to have been differentially manifested in surface form well into early Modern English times, as shown in (99) and (100), they are currently manifested prosodically only. Note that in present-day English either particle here, which is superficially invariant, is normally more prominent prosodically when it follows the lighter verb *to be* than when it does the heavier verb *to come* or *to go*. This particle alternation, which also is apparently osmotically triggered, is thus clearly audible in present-day English as well, though not formally “visible” (unlike in Middle English or early Modern English).

Underlying the prosodic difference here, there may be a subtle difference in semantic mass/weight between the two tokens of either particle under discussion. Either particle may be heavier in meaning when it combines with the lighter *to be* than when it does with the heavier *to come* or *to go*. This, if correct, is arguably because the particle has less osmotic concession of meaning to make to the lighter *to be* than it does to the heavier *to come* or *to go*. We will have more to say about this kind of osmotically-triggered semantic variation in Section 5.

The two different manifestations of the preposition *of* in early Modern

English, as illustrated by the first member of the following paraphrase pair, may also be ascribable to differentially applied osmotic pressure.

- (101) a. Which *of* my ships art thou master *off*?
 b. Which *of* my ships are you master *of*?

The first sentence here is the Shakespearean version of the second sentence from present-day English. Note that in Shakespearean English the preposition in question here was manifested as *of* with a single *f* in the presence of an immediately following complement and as *off* with a double *f* in the absence of such a complement. Thus the preposition assumed the lighter/shorter form *of* when it had to share prepositional-phrase space with an immediately following complement and the heavier/longer form *off* when it did not have to do so. It may be noted here that *off* was heavier than its alternant *of* not just in spelling but also in pronunciation so that the alternation here may be cited in support of orthographic/phonological osmosis as well.

Just as is the case with the particles *in* and *out*, both tokens of the preposition *of* here are presently manifested identically in orthophonemic form, but differentially in prosody. The first token of *of* in (101b), which corresponds to the Shakespearean *of*, is prosodically less prominent than the second token of the same in (101b), which corresponds to the Shakespearean *off*. We may point out here that similarly explainable is the prosodic behavior of other prepositions with respect to the alternation between their stranded and non-stranded versions with the former normally outweighing the latter (semantically and thus also) prosodically.

Osmotic pressure also apparently sheds light on why *to behave*, but not its negative counterpart *to misbehave*, may allow a direct object in the form of a reflexive pronoun, as illustrated in the sentence pair below.

- (102) a. If you *behave* (*yourself*), I'll let you stay up to watch the movie.
 b. If you *misbehave* (**yourself*), I won't let you stay up to watch the movie.

Note here that *to behave* is considerably lighter in meaning and form than is *to misbehave* and thus the latter, unlike the former, arguably cannot afford the room to osmotically accommodate the reflexive object

yourself. It is interesting that the paraphrase *to behave badly* for *to misbehave* also disallows a reflexive object here, so that **to behave oneself badly* is just as unacceptable as **to misbehave oneself*. For *to behave badly* (as well as *to behave kindly*, etc.) is just as heavy as *to misbehave* so that it, too, cannot afford osmotic room for a reflexive object.

Our discussion here may help provide an enlightened explanation for the optional absence of the indefinite article *a* from *to catch (a) cold*, as opposed to its obligatory presence in *to have a cold*, as can be seen from the two paraphrase pairs below.

- (103) a. He *caught a cold*.
b. He *caught cold*.

- (104) a. He *had a cold*.
b. **He had cold*.

Note that the dynamic *to catch* is slightly heavier (in meaning) than the stative *to have* here and that the latter, but not the former, always requires the surface support of *a* for the object noun phrase *cold*. Arguably osmosis-conditioned, this difference in indefinite-article usage may be designed to put on a rough formal par the two “equivalent” verb phrases here, i.e. the *catch*-initial verb phrase and its *have*-initial equivalent.

Osmotic pressure can also be invoked in explaining why the indefinite article figures in the first member of each paraphrase pair below, but not in the second.

- (105) a. It *is a great pleasure/delight/honor* for me to do this for you.
b. It *gives me great pleasure/delight/honor* to do this for you.

Suppose that the indefinite article *a* were not used following the verb *to be* in the first sentence here. Since the stative verb *to be* is less heavy (in meaning and form) than the dynamic verb *to give*, the italicized verb phrase would then end up being so much lighter in the first sentence than in the second. This would apparently disturb formal equilibrium between the two cognately equivalent phrases in question here.

Thus viewed, the indefinite article *a* here arguably serves to boost the formal mass/space of the *is*-initial verb phrase up to the level of the *gives*-initial verb phrase. If this is correct, then the indefinite article *a* here may function as osmotically-triggered padding of sorts that helps the two cognately equivalent verb phrases in question to attain a balance of formal mass/space. Thus the indefinite article may be more of an adjuster of formal mass/space than a numerical counter here.

It is worth noting in this connection that the choice between count and non-count usages of one and the same noun here (e.g. *pleasure*) is apparently determined by the relative weight of the verb with which it occurs. If this is indeed the case, then the count/non-count distinction in English grammar may be something of a misnomer when it comes to cases such as those involved in (103)-(105) above. For it does not seem to have too much to do with the count/non-count distinction in the normal logical sense of the term.

An essentially identical rationalization is applicable to the presence of the indefinite article *a(n)* in the first member of each pair below, as opposed to its absence from the second member of the same pair.

- (106) a. The conference *was a success/failure*.
 b. The conference *ended in success/failure*.
- (107) a. Their last campaign *was a great disappointment*.
 b. Their last campaign *ended in great disappointment*.
- (108) a. She *was a disgrace* to womankind.
 b. She *brought disgrace* upon womankind.
- (109) a. Reading *is always a joy* to me.
 b. Reading *always brings joy* to me.
- (110) a. The news of his safe return *was a great relief*.
 b. The news of his safe return *brought great relief*.

We are assuming here that the verb *to be* is lighter than the verb *to end in* or *to bring*.

The same explanation applies to the indefinite article that figures in the first member of either paraphrase pair below, but not in the second

member of the same pair.

- (111) a. The source of the gunshots still *remains a mystery*.
 b. The source of the gunshots *is still shrouded in mystery*.
- (112) a. Dr. Johnson *is a credit* to his profession.
 b. Dr. Johnson *does his profession credit*.

Again, we are assuming here that *to remain* and *to be* are lighter than *to be shrouded in* and *to do* respectively.

It is interesting to speculate at this point that an osmotically-triggered search for formal parity between cognately equivalent phrases apparently underlies the subjective-complement alternation illustrated by paraphrase pairs such as the following.

- (113) a. His health is *a concern* to all of us.
 b. His health is *of concern* to all of us.
- (114) a. This item is *a good value*.
 b. This item is *of good value*.
- (115) a. Those picture dictionaries have been *a great help* to me.
 b. Those picture dictionaries have been *of great help* to me.

Formal parity is arguably achieved and maintained between the two alternative subjective complements in each pair here through the mutually exclusive distribution of the indefinite article *a* and the preposition *of*.

It is not just the indefinite article that may get triggered by osmotic pressure. The definite article may get similarly triggered, as can be seen from an examination of the following paraphrase pairs.

- (116) a. Georgetown was ahead 47:38 at *halftime*.
 b. Georgetown was ahead 47:38 at *the half*.
- (117) a. He's down with *influenza*.
 b. He's down with *the flu*.

- (118) a. He's got *delirium tremens*.
 b. He's got *the DTs*.
- (119) a. He was at *home plate* for his first at-bat.
 b. He was at *the plate* for his first at-bat.

The definite article here arguably serves to fill in the vacuum left by the clipping of *halftime* to *half* in (116), of *influenza* to *flu* in (117), of *delirium tremens* to *DTs* in (118) and of *home plate* to *plate* in (119). Admittedly, *the* in *the flu* is optional for some speakers, for whom (117) may lose some of its validity as a supportive example for our claim here. Note that the deletion-compensating *the* here is arguably on the anaphoric side since it may be thought of as referring back to the portion deleted from the underlying form in question.

Incidentally, *the pox* (in the sense of "syphilis") may stem via clipping from either *chickenpox* or *smallpox* or both. If this is correct, then the definite article *the* here also arguably serves to osmotically compensate for the portion deleted in the derivation in question, i.e. for either *chicken-* or *small-* or both.

Similarly rationalizable is the italicized definite article used in the second member of either paraphrase pair below.

- (120) a. They live not too far from *Veterans Stadium*.
 b. They live not too far from *the Vet*.
- (121) a. The Cleveland Indians play all their home games at *Jacobs Field*.
 b. The Cleveland Indians play all their home games at *the Jake*.

Apparently of similar genesis is the optional *the* in the second member of the following paraphrase pair.

- (122) a. Katie Couric was chatting away on *television*.
 b. Katie Couric was chatting away on *(the) TV*.

Another interesting example of a clipping-triggered, deletion-compensating definite article is afforded by the following quote from *The New American Bible* (Luke 18:14).

- (123) For *everyone* who exalts himself will be humbled, and *the one* who humbles himself will be exalted.

The definite article in *the one*, which replaces *everyone* here, is apparently a compensatory fill-in for *every-* in *everyone*. Incidentally, that *the one* replaces *everyone* here is, in fact, indirectly supported by *The Good News Bible*, which renders (123) as (124) below.

- (124) For *everyone* who makes himself great will be humbled, and *everyone* who humbles himself will be made great.

Apparently explainable along similar lines is the definite article in *the masks* in the following quote from the October 31, 2001 issue of *USA TODAY*, a mass-circulation U.S. daily newspaper.

- (125) *Safety masks* can provide some protection against anthrax, but no one can say for certain how much. Still, orders for *the masks* are way up.

Since *the masks* here is a generic “noun” that replaces the compound noun *Safety masks*, *the* in the former is arguably a compensatory fill-in for *Safety* in the latter.

It is interesting to note in this connection that osmotic pressure arguably also underlies the replacement of *those who are* and *that which is* by *the* in paraphrase pairs such as the following.

- (126) a. There's one law for *those who are rich* and another for *those who are poor*.
 b. There's one law for *the rich* and another for *the poor*.
- (127) a. From *that which is sublime* to *that which is ridiculous* is but a step.
 b. From *the sublime* to *the ridiculous* is but a step.

Apparently explainable along similar lines is the italicized definite article that figures in the second member of each paraphrase pair below.

- (128) a. Lots of things happen on *Capitol Hill*.
b. Lots of things happen on *the Hill*.
- (129) a. The president was on his way to *Cape Canaveral*.
b. The president was on his way to *the Cape*.
- (130) a. Prospects are bright for talks between North Korea and *South Korea*.
b. Prospects are bright for talks between North Korea and *the South*.
- (131) a. *Vatican Hill/City/Palace* is a popular tourist destination.
b. *The Vatican* is a popular tourist destination.
- (132) a. Gordon is the favorite in today's NASCAR race at *Watkins Glen*(, *New York*).
b. Gordon is the favorite in today's NASCAR race at *the Glen*.

Osmotic pressure also is apparently involved in the use of the italicized *the* in the second member of each of the following paraphrase pairs.

- (133) a. *To hell* with the rules, I'm going to do it anyway.
b. *The hell* with the rules, I'm going to do it anyway.
- (134) a. Speaker A: I won't tell, I swear.
Speaker B: *Like hell* you won't.
b. Speaker A: I won't tell, I swear.
Speaker B: *The hell* you won't.

Incidentally, it is tempting to speculate here that *the hell* in sentences like *What the hell are you talking about?* derives from *in hell*. If this is correct, then we may say that *the* is a compensatory fill-in for *in* here also. However, the problem here is that we often encounter forms like *What in the hell are you talking about?*, which seems to militate against our speculation.

In further support of our contention here, we may cite the following paraphrase pair, the second member of which is from *The New Jerusalem Bible (Mark 10:31)*.

- (135) a. Many who are first will be last, and *many who are last* will be first.
 b. Many who are first will be last, and *the last*, first.

The definite article in *the last* in (135b) is arguably an osmotically-triggered compensatory trace for *many who are*, i.e. the portion that gets deleted from *many who are last* in (135a). It is worth noting that (135b) is rendered in *The Good News Bible* as *Many who are now first will be last, and many who are now last will be first*, which is (practically) identical to (135a) above.

Likewise, the underlined definite article in each paraphrase pair below may be an osmotically-triggered compensatory trace for the underlined plural suffix *-s* (or vice versa).

- (136) a. President Bush opposes any cloning for *purpose~~s~~* of human replication.
 b. President Bush opposes any cloning for *the purpose* of human replication.
- (137) a. We kept searching in *hope~~s~~* of finding them alive.
 b. We kept searching in *the hope* of finding them alive.

Similarly interpretable may be the complementary distribution of the plural suffix *-s* and the article, either definite or indefinite, in generic noun phrases such as those in italics in the paraphrase set below.

- (138) a. *Tiger~~s~~* are dangerous animals.
 b. *The tiger* is a dangerous animal.
 a. *A tiger* is a dangerous animal.

It is interesting that osmotic pressure should arguably underlie the mutually exclusive distribution of the plural suffix and the article here, which apparently is not confined to the examples given here but quite widespread in English. Incidentally, given the synonymy between the alternants in question here, i.e. between *Tigers* and *The/A tiger*, the semantic boundary between singular and plural may be fuzzy or even dubious when it comes to generic noun phrases. Assuming in this connection that *a large number of people* and *large number~~s~~ of people*

are paraphrases of each other, we can rationalize the *a/-s* alternation here along identical lines also.

It is interesting that the (definite) article is often mass/space-sensitive in the other direction as well. Rather than being tacked on as padding or ballast, it is often discarded very much like jetsam that gets cast overboard to help reduce the mass/space of an overloaded or overcrowded form. For example, (the singular form of) a count noun, which normally occurs with a preceding article, may not allow such an article when it is reinforced by the prefix *mid-*, as can be seen from pairs of paraphrase sentences such as the following.

- (139) a. Nixon had to resign in *the middle of the term*.
 b. Nixon had to resign in *midterm*.
- (140) a. We had a meeting scheduled for *the middle of the week*.
 b. We had a meeting scheduled for *midweek*.
- (141) a. They crossed the Rockies in *the middle of (the) winter*.
 b. They crossed the Rockies in *midwinter*.
- (142) a. Norman was forced to quit in *the middle of the match*.
 b. Norman was forced to quit in *midmatch*.
- (143) a. Don't change horses in *the middle of the stream*.
 b. Don't change horses in *midstream*.

We may argue here that *mid-* renders a little too heavy the noun to which it is prefixed with the result that there is hardly any room left for predetermination by an article, that is, within the limited confines of the noun phrase in which it occurs.

The prefixation here may be thought of as creating something of a super-compound noun that fills up a whole noun-phrase slot. The noun phrase of which this super-compound noun is the head is evidently so expanded and thus dense in mass/space that it can hardly afford any osmotic room for predetermination by an overtly manifested definite article. It may be that this super-noun is something of a linguistic black hole into which such underlying matter as the definite article here gets osmotically absorbed so thoroughly that it becomes impervious to sight.

At any rate, the super-noun in question here is apparently very much like a non-count (or proper) noun in formal behavior in that normally it also is non-count and does not allow predetermination by an article.

Assuming that the second member of each paraphrase pair in (139)-(143) above derives from the first member of the same pair, the weight gain of the *mid*-initial noun (phrase) here is arguably the result of “downsizing” a two-phrase nominal to a one-phrase nominal. This downsizing apparently involves squeezing the phrase *the middle of* into the noun phrase that it underlyingly pre-modifies. In fact, the matrix noun phrase here can afford so little extra osmotic room that not just the *the* - *-dle of* of the insert phrase *the middle of* but also the *the* of the matrix noun phrase ends up getting squeezed or zeroed out.

It is interesting to note in this connection that *year's*, as in *year's end*, is evidently a quasi-prefixal reinforcer on a formal par with *mid-*, as can be seen from an analysis of the following paraphrase pair.

- (144) a. The report was released at *the end of the year*.
 b. The report was released at *year's end*.

Assuming that the second member of the above pair derives from the first, we may argue that the insert *of the year* is squeezed into the matrix *the end*, in the course of which not just *of the* in the insert but also *the* in the matrix disappears under osmotic pressure. Thus *year's end* apparently dispenses with predetermination by the definite article for basically the same reason that, say, *midyear* does as in *I like the warm weather in midyear*.

Note here that *year's end* is treated as a mono-lexemic super-compound noun although it is orthographically bi-lexemic. Given plenty of highly similar (mono-lexemic) compounds such as *linesman*, *woodsman* and *landsmen* with the “word-medial” suffix *-s*, it does not seem too farfetched to regard *year's end* as a (mono-lexemic) super-compound.

Apparently lending additional credence to the quasi-prefixal status of *year's* in *year's end* in (144) above is the productivity of expressions of like type in English, as can be attested to by an abundance of paraphrase pairs such as the following

- (145) a. Everything was back to normal by *the end of the month*.
 b. Everything was back to normal by *month's end*.

- (146) a. He went on the disabled list near *the end of the season*.
 b. He went on the disabled list near *season's end*.
- (147) a. I'll just be teaching here till *the end of the semester*.
 b. I'll just be teaching here till *semester's end*.

It is perhaps in order here to point out that the super-compound *year's end*, for one, may sometimes get downgraded to the regular compound *year-end*, which is often preceded by the definite article, as can be seen from the following paraphrase pair.

- (148) a. Sales will go up toward *year's end*.
 b. Sales will go up toward *the year-end*.

Assuming that *year-end* derives from *year's end*, we may argue that the definite article comes back to life in (148b) to make up for the mass/space lost to the deletion of the genitive suffix *'s*, arguably under osmotic pressure. An example such as the following paraphrase pair seems to militate against this argument, however.

- (149) a. Executives hope to raise share prices 4% by *year's end*.
 b. Executives hope to raise share prices 4% by *year-end*.

According to *Longman Advanced American Dictionary*, from which (149b) is taken, *year-end* is a non-count noun. If it is indeed a bona fide non-count noun, *year-end* should normally occur without an article, as in (149b) above. In fact, this analysis of *Longman's* seems to be borne out by the following examples taken from "2002 Market Outlook" in *USA TODAY* (January 2, 2002).

- (150) a. By *year-end*, he expects the market to be up 5% from where it is now.
 b. Assuming a meaningfully higher market at *year-end* is a very aggressive view.

Given (149) and (150) alongside of (148), we should probably assume that *year-end* is on the "count/non-count" borderline, thus only optionally allowing predetermination by the definite article *the*.

However, examples of *year-end* in most other dictionaries seem to occur with the definite article, which may help validate our comment about the coming-back-to-life of *the* in connection with (148) above. Furthermore, the derivation illustrated by the italicized portions of the following paraphrase set may lend some analogical force to our comment in question here.

- (151) a. We'll be back by *the end of the week*.
 b. We'll be back by *week's end*.
 c. We'll be back by *the weekend*. / ?We'll be back by *weekend*.

The analogy argument presented here may not be quite as convincing as it might at first appear to be, however. It may be that *weekend* has already become a regular compound and is no longer a super-compound, whereas in the case of *year-end* the change from super- to regular compound is a "work in progress" and thus not yet completed. Thus *week-* in *weekend* may no longer be an (article-displacing) quasi-prefixal reinforcer, which *year-* in *year-end* may optionally be.

We may note in this connection that *New Year's* is also a quasi-prefixal reinforcer on a par with *year's*, as can be seen from the italicized parts of the following sentences.

- (152) a. What are your plans for *New Year's Day*?
 b. They were married on *New Year's Eve*.
 c. We plan to get together on *New Year's morning*.

Season's in *season's greetings* is a quasi-prefixal reinforcer of a similar nature, as can be seen from the following paraphrase pair.

- (153) a. I was busy sending out (*the*) *greetings of the season*.
 b. I was busy sending out *season's greetings*.

We may also argue along similar lines that the name of a day of the week premodifying such temporal nouns as *morning*, *afternoon* and *night* is a quasi-prefixal reinforcer on the order of *mid-*, *year's* or the like, in view of paraphrase pairs such as the following

- (154) a. I'll see you on *Sunday in the morning*. / (?)I'll see you on *the morning of Sunday*.
 b. I'll see you on *Sunday morning*.
- (155) a. I came on *Monday in the afternoon*. / (?)I came on *the afternoon of Monday*.
 b. I came on *Monday afternoon*.
- (156) a. The concert is scheduled for *Saturday in the evening*. /
 (?)The concert is scheduled for *the evening of Saturday*.
 b. The concert is scheduled for *Saturday evening*.

Incidentally, the name of the day that figures in any pair here may refer to a day other than a day of the week so that, for example, *Labor Day* may replace *Sunday* in (154) above.

It is interesting to note at this point that the ordinal numeral *Twelfth* in the three italicized portions of the sentence below is evidently another (article-displacing) quasi-prefixal reinforcer of the sort under discussion here.

- (157) *Twelfth Tide* comprises *Twelfth Night* and *Twelfth Day*.

Our discussion here sheds light on the absence of the definite article from names of holidays such as those cited in the sentences below.

- (158) a. The game is scheduled for *Christmas Day*.
 b. They were married on *Independence Day*.
 c. *Armed Forces Day* honors all branches of the military.

Christmas, *Independence* and *Armed Forces* here are arguably quasi-prefixal reinforcers that combine with *Day* to form the three italicized super-compound nouns given in the sentence set above. The noun phrases of which these super-nouns are heads are arguably so dense in mass/space that they can hardly afford any osmotic room for predetermination by an article, with which *Day/day* normally occurs elsewhere. Compound names of days of significance other than official holidays normally behave in exactly the same way in that they also dispense with the definite article, as in *graduation day*, *homecoming*

day, opening day, media day, etc. Incidentally, *day* may be replaced by *eve, morning, afternoon, evening* and *night* in all our examples here.

It may be in order here to note that the “zeroing out” of the (definite) article under discussion may sometimes be optional, as can be seen from the two alternate versions of one and the same Biblical message (*Matthew 12:41*) quoted below.

- (159) a. On *Judgement Day* the men of Nineveh will appear against this generation and they will be its condemnation --.
 b. On *the Judgment Day* the people of Nineveh will stand up and accuse you --.

The first version here, which comes from *The New Jerusalem Bible*, dispenses with *the*; the second version, which comes from *Good News Bible*, does not. It may be that *Judg(e)ment Day* is a super-compound noun for some people and an ordinary noun phrase (or a regular compound noun) for others.

We may contend at this point that what precedes *-day* in the name of a day of the week, e.g. *Satur-* in *Saturday*, is also a quasi-prefixal reinforcer of a similar nature. If correct, this should help explain why the name of a day of the week is normally not preceded by an article. By the same token, *yester-* in *yesterday/yesteryear* and *to-* in *today/tonight/tomorrow* also may be quasi-prefixal reinforcers of like kind that serve to block articles from in front of the nouns in which they occur.

Quasi-prefixal reinforcers of the sort under discussion here may be followed by temporal nouns other than *day, morning, afternoon* or the like, as can be seen from examples such as the following.

- (160) a. February is *black history month*.
 b. Was *International Geophysical Year* 1959 or 1960?
 c. When's *happy hour* here?
 d. Versace's haut couture show Saturday kicked off *fashion week* in Paris.

It is worth noting in this connection that *Longman Advanced American Dictionary* lists *rush hour* as ambivalent between count and non-count status. It may be that *rush hour* is either a noun phrase in the process of getting compacted into a super-compound noun or a super-

compound noun in the process of getting downgraded to a regular compound noun. It arguably allows predetermination by an article when it is perceived as either a noun phrase or a regular compound noun, while it disallows such predetermination when it is perceived as a super-compound noun. This may have led the *Longman* lexicographers to classify *rush hour* as ambivalent between count and non-count.

Used as premodifiers (of certain frequently-used temporal nouns), *early* and *late* are often very much like quasi-prefixal reinforcers of the sort under discussion here. In support of this contention, we may point out that they tend to dislodge the definite article from expressions in which they premodify *morning*, *afternoon* and other such nouns (designating parts of a day). Let us take the paraphrase pair below, for example.

- (161) a. They usually take a walk *early/late in the morning*.
 b. They usually take a walk *in early/late morning*.

Let us now turn our attention to the fact that the names of the four seasons may or may not be preceded by the definite article, as can be seen from the italicized noun phrases in the following sentences.

- (162) a. They go to Hawaii in *(the) summer*.
 b. Does it snow much in *(the) winter*.
 c. The festival is usually held in *(the) spring*.

It is interesting that the optionality of the definite article in expressions like the italicized ones in (162) drops significantly when the names of the seasons are premodified by such adjectives as *early* and *late*. Thus expressions like those in italics below more often than not dispense with the definite article, especially when the reference is generic.

- (163) a. They used to come and see us in *early spring*.
 b. We get lots of snow in *late winter*.

The premodifiers *early* and *late* here may add quite a bit to the formal mass/space of the names of the seasons in question much as do the quasi-prefixal reinforcers we have just considered, so that the noun phrases in question may not have too much osmotic room left for the definite article.

Incidentally, *all* in such temporal adverbials as those illustrated in the following sentences is evidently reinforcing material of a similar nature, which may help explain why they may not contain the definite article.

- (164) a. They worked there *all day (long)*.
 b. They stayed there *all summer (long)*.
 b. It's been raining *all week (long)*.

In this connection, let us compare the two sentences below, each of which is a paraphrase of the other.

- (165) a. It's warm enough to swim *all year round*.
 b. It's warm enough to swim *all the year round*.

Although *all year round* and *all the year round* are interchangeable here, we probably would feel more comfortable with the former than with the latter. It may be that the adverbial in question is so reinforced (by *all - round*) that there is not too much room left for *the* in front of *year* here. Apparently rationalizable in a similar fashion is the free variation between *at this time of year* and *at this time of the year*, of which the former is probably more common than the latter (unless *year* here has a distinctly specific reference).

Next and *last* in sentences such as those below may also be quasi-prefixal reinforcers of a similar nature that serve to dislodge a predetermining article.

- (166) a. See you *next week*.
 b. We were in London *last year*.

Note at this point that *next* and *last* as well as *all - (long)* and *all - round* are apparently such potent reinforcers that they displace not just articles but also prepositions. This is apparently why the temporal expressions they quasi-prefixally reinforce are not normally preceded by either prepositions or articles.

Our discussion from (139) to (166) has focused on temporal expressions from which the article is "bounced" when they are reinforced by quasi-prefixal elements. We can just as easily support our thesis here, however, by citing examples of non-temporal expressions such as those

that figure in the derivation of the second member of either paraphrase pair below from the first member of the same pair.

- (167) a. Seattle is the largest city in *the state of Washington*.
 b. Seattle is the largest city in *Washington state*.
- (168) a. The 2002 World Economic Forum was held in *the City of New York*.
 b. The 2002 World Economic Forum was held in *New York City*.

We can argue here that the definite article is zeroed out when the bi-phrasal *the state of Washington* and *the City of New York* are turned into “mono-phrasal” super-compound names, i.e. into *Washington state* and *New York City* respectively. We would argue that *Washington* and *New York* in *Washington state* and *New York City* are quasi-prefixal reinforcers that add so much to the formal mass/space of the compounds in question here that there is no osmotic room left for overt predetermination by the definite article. It goes without saying that an identical explanation applies to the absence of the definite article from such similar place names as *Atlantic City*, *Mexico City* and *Guatemala City*. It may be observed in this connection that *the Evergreen State*, the descriptive nickname for *Washington state*, begins with the definite article *the* because it is not a super-compound noun but either a regular compound noun or an ordinary noun phrase.

An identical argumentation is apparently applicable to the deletion of the definite article in the derivation of the second member of each paraphrase pair below from the first member of the same pair.

- (169) a. He’s teaching at *the University of Cambridge*.
 b. He’s teaching at *Cambridge University*.
- (170) a. She majors in *the history of Korea*.
 b. She majors in *Korean history*.

At this point, we may invoke the notion of quasi-prefixal reinforcer or of super-compound noun in explaining the absence of the definite article from such compound place names as the italicized ones in the sentences below.

- (171) a. As usual, *Yankee Stadium* was packed with baseball fans.
 b. *Wrigley Field* was not exactly his favorite baseball park.
 c. The Dodgers have so far been out of luck at *Candlestick Park*.
 d. The New York Knicks play their home games at *Madison Square Garden*.
 e. Practice opens Friday at *Daytona International Speedway*.

Although phrasal in appearance, the names in italics here are all super-compound nouns, in which the underlined parts are quasi-prefixal reinforcers that arguably eventuate in the non-use of the definite article here.

Needless to say, the italicized compound names in the following sentences are amenable to an identical explanation.

- (172) a. He has an office on *Fifth Avenue*.
 b. *Wall Street* is the major financial center of the United States.
 c. They used to live on *Wilson Boulevard*.
 d. *JFK International Airport* is not the only airport serving New York City.
 e. He runs a snowboarding camp for kids at *Stratton Mountain*.
 f. I prefer *Alamoana Beach* to *Waikiki Beach*.
 g. The marchers were supposed to converge on *Windsor Castle*.

The type of structure under discussion here seems to be quite widespread in English, as can be attested to by additional examples such as the following.

- (173) a. *Capitol Hill* is the small hill in Washington, D.C., on which the Capitol stands.
 b. Our flat is not too far from *Buckingham Palace*.
 c. *Notre Dame Cathedral* is a popular tourist destination.
 d. *Scotland Yard* was formerly the site of the London police headquarters.
 e. *Temple Square* is Utah's #1 tourist attraction.
 f. *Communism Peak* is in Tadjikistan.
 g. *Clarence House* is the Queen Mother's residence in London.

We can add to this list almost ad infinitum by citing examples such as the following: *Sherwood Forest*, *Yellowstone National Park*, *Westminster Abbey*, *Westminster Hall*, *Carnton Plantation*, *City Hall*, *City Hall Plaza*, *Orange County*, *Pleasant Grove*, *Flathead Lake*, *Coconut Island*, *Emerald Isle*, *Camden Yard*, *Sea Island Golf Club*, *America West Arena*, *Pauley Pavilion*, *Immaculate Conception Church*, *Andrews Air Force Base*, *Hunter Airfield*, *Jackson Memorial Hospital*, *Arlington National Cemetery*, *Alexandria Township*, *Hunan Province*, *Hokkaido Prefecture*, etc.

Incidentally, names of the sort under consideration here may sometimes be preceded by an article when they are perceived as noun phrases that have not yet evolved into super-compound nouns. Thus there seems to have been a time when London's *Green Park*, for one was referred to as *the Green Park*, that is, before the name gelled into the super-compound noun *Green Park*.

It is interesting to note in this connection that *horseback* is apparently a super-compound noun of the sort under discussion, as can be seen from the following paraphrase pair.

- (174) a. The sheriff arrived on *the back of a horse*.
 b. The sheriff arrived on *horseback*.

Unlike the non-temporal super-compound nouns we have considered prior thereto, *horseback* here is not proper-nounish but common-nounish. Common-nounish super-compounds also seem to be fairly common in English, as can be seen from an abundance of such examples as those italicized in the sentences below.

- (175) a. Teresa Earnhardt was in *victory lane* to congratulate Dale Earnhardt, Jr.
 b. He's on *death row* for murdering eight people.
 c. The two teams are tied for *first place*.
 d. I once won *first prize* in the contest.
 e. They travel either *economy class* or *business class*, never *first class*.
 f. *Tennis elbow* is a serious medical problem for lots of my colleagues.
 g. He subscribes to *TIME magazine*.

- h. GM is bleeding *market share* and Ford is poised to overtake it.
- i. *Baseball season* opens in seven cities Friday.

Admittedly, nominal compounds of the sort illustrated here sometimes seem to fluctuate between (super-compound) noun and noun phrase. *Season*-headed nominals such as *baseball season* (as in (175i) above) may be a case in point. Thus *baseball season* (in its generic sense) may or may not require article predetermination although it normally seems to dispense therewith.

Another interesting case of a common-nounish super-compound of the sort under consideration here is an array of *fever*-headed nominals such as *hay fever* illustrated in (176b) below.

- (176) a. The girl was coming down with *a fever*.
- b. The girl was coming down with *hay fever*.

Note that *fever* in the first sentence here is a regular common noun so that it is preceded by the article *a*. Note further that *hay fever* in the second sentence above is a super-compound noun headed by the common noun *fever* with *hay* functioning as a quasi-prefixal reinforcer of the sort under discussion here. This is evidently why *hay fever* is not preceded by an article although its head *fever* itself is a common noun. Exactly the same account applies to the absence of article predetermination from the names of not just such other physical fevers as *scarlet fever*, *typhoid fever*, *rheumatic fever*, etc. but also such figurative fevers as *spring fever*, *cabin fever*, *travel fever*, *diet fever*, etc.

The quasi-prefixal reinforcer in question here may sometimes be sense-selective in that it is keyed to the particular sense of the "stem" noun to which it is prefixed or the idiomaticity of the nominal in which it occurs. Let us take the following pair of sentences (from *Longman Advanced American Dictionary*), for example, focusing on their subject noun phrases.

- (177) a. *Herniated disk syndrome* is more common in males.
- b. "*The underdog syndrome*" is a belief that things are beyond your control.

Note that *syndrome* refers to a physical condition in the first member of the pair here, as opposed to a non-physical condition in the second. Note also that *herniated disk syndrome* refers to a conceptually more “institutionalized” condition than does *underdog syndrome* so that the former may be more of an idiomatic super-compound noun and less of a syntactic phrase than the latter. Thus *herniated disk*, but not *underdog*, is a quasi-prefixal reinforcer of the sort under discussion here, which is why the definite article is absent from in front of *herniated disk syndrome*, but not from in front of *underdog syndrome*. Exactly the same account applies to the absence of the article from in front of *premenstrual syndrome*, *acquired immunodeficiency syndrome*, etc., but not from in front of *it-won't-affect-me syndrome*, *feast-or-famine syndrome*, etc.

Our point seems to be well taken in light of further examples such as the one afforded by the italicized noun phrases in the following pair of sentences.

- (178) a. We are going skiing over *spring break*.
 b. Let's take *a coffee break*.

Note that the two tokens of *break* here are differentiated sense-wise in that the first token refers to a formal vacation and the second to a (short) informal period of rest. Besides, *spring break* is conceptually better institutionalized or conventionalized and is thus more of an idiomatic super-compound noun and less of a syntactic phrase than is *coffee break*, so that *spring* here, but not *coffee*, is a quasi-prefixal reinforcer of the article-displacing sort. Incidentally, some lexicographers list *spring break* as a count noun, which may be a somewhat controversial position to take in the light of the fact that it may often be non-count, as in (178a) above, although it may just as well be count. Needless to say, we can explain in exactly the same way why *Christmas break*, *Thanksgiving break*, etc. normally do away with predetermination by an article, whereas *five-minute break*, *exercise break*, etc. may not.

We may cite the following additional example to make the point that quasi-prefixal reinforcers are often sense-selective.

- (179) a. I grew up in *Amish country*.
 b. I grew up in *the/an Asian country*.

Note that the italicized premodifier here is an (article-displacing) quasi-prefixal reinforcer when *country* means “land characterized by a unique trait,” but not when it simply means “land where the people of a nation live.” Note further that *Amish country* is perhaps better institutionalized or conventionalized than *Asian country* in that the former normally refers to a definitionally more unique socio-cultural entity than does the latter. Thus *Amish country* is more of an idiomatic super-compound noun and less of a syntactic phrase than *Asian country*, so that *Amish*, but not *Asian*, qualifies as an (article-displacing) quasi-prefixal reinforcer.

The relative lexicalization of a nominal compound in question may also have to be factored into our discussion of whether or not a given premodifier in it qualifies as a quasi-prefixal reinforcer of the sort under discussion here. Let us consider the following pair of sentences with reference to the italicized compound nouns.

- (180) a. He can play either *the infield* or *the outfield*.
 b. He'd rather play *left field* than either *center field* or *right field*.

According to *Random House Webster's Dictionary*, the italicized compound nouns in the first sentence here are about two centuries older than those in the second, the former having come into the language sometime between 1600 and 1640 and the latter in the 1850s. Thus the first, premodifying constituent of the compound noun here may have had more time to merge with the second, premodified constituent of the same noun in the first sentence above than in the second. In other words, the italicized nominal compounds are arguably further along in lexicalization in *infield* and *outfield* than in *left field*, *center field* and *right field*. Incidentally, this difference is manifested superficially by the absence of orthographic space between the two immediate constituents in question in *infield* and *outfield*, as opposed to the presence thereof in *left field*, *center field* and *right field*.

The two immediate constituents in question here may thus have been in closer construction with each other and have made more osmotic concession of mass/space to each other in the first sentence than in the second. As a result, the nominal compound that they make up may have become more of a regular compound (and less of a super-compound) in

(180a) and the other way around in (180b). If this is correct, then there should be less osmotic room for the definite article in front of the compound in (180b) than in front of (180a). This is arguably why we end up with *the* in front of *infield* and *outfield*, but not in front of *left field*, *center field* and *right field*.

Alternatively, it may be that *in-* and *out-* here are inherently lighter in formal/semantic mass/space than are *left*, *center* and *right* so that the latter, but not the former may qualify as (article-displacing) quasi-prefixal reinforcers. Or, the use of the definite article in front of *infield* or *outfield*, as opposed to *left field*, *center field* or *right field*, may be a linguistic concession to the former occupying a larger stretch of playing field than does the latter.

Whatever may be the case, the definite article may be absent from the italicized nominals in the examples below for apparently the same reason it is absent from the italicized nominals in (180b) above.

- (181) a. John wants to play *first base*.
 b. I got picked off trying to steal *second base*.
 c. He was safe at *third base*.
 d. I used to play *short stop*.
 e. As usual, the umpire was behind *home plate*.

Incidentally, *Longman Advanced American Dictionary* lists *first base* and *center field* as count nouns, and *left field* and *right field* as singular nouns. Regarding *first base* and *center field* here as count nouns may be on the controversial side. For they often seem to be more like non-count nouns than count nouns on the surface at least, in that they are normally not preceded by an article. Regarding *center field* as a count noun, while treating *left field* and *right field* as singular nouns, also seems to be inconsistent at best. For, among other things, they all behave identically in that they normally do not allow predetermination by an article.

Quasi-prefixal reinforcers appear to displace not just the definite article but also the indefinite article, as can be seen from an examination of the following paraphrase pair.

- (182) a. He predicted the outcome of the contest with *an accuracy that was remarkable*.
 b. He predicted the outcome of the contest with *(a) remarkable accuracy*.

Note here that the quasi-prefixal reinforcer *remarkable* more often than not displaces the indefinite article *a(n)* so that the second sentence above probably terminates more often with *remarkable accuracy* than with *a remarkable accuracy*. The indefinite article here tends to get pushed out, arguably under osmotic pressure, when the bi-phrasal *an accuracy that was remarkable* is compacted into the mono-phrasal *(a) remarkable accuracy*.

Parenthetically, the idiomatic *the dead of*, as in *the dead of winter* and *the dead of night*, may also be regarded as something of a quasi-prefixal reinforcer of the sort under discussion here. This may help explain why *winter* and *night* here are not (immediately) preceded by an article, either definite or indefinite. Among reinforcers of a similar nature may be *kind of*, *sort of*, *breed of*, *make of* and *office/job of*, as in *another kind of flower*, *a new sort of computer*, *a good breed of dog*, *a different make of computer* and *the office/job of president*.

The article-suppression phenomenon under observation here is discussed in greater detail (along with other kinds of structural suppression) in Park in preparation ("Osmotic Depletion in Structural Condensation"), which will be focusing on examples such as the following.

- (183) a. Flags were flown at *half-mast*.

We will be arguing that *half* here is a quasi-prefixal reinforcer that bounces the definite article from the compound noun headed by *mast* in the (idiomatic) super-compound *.at half-mast*. We also will be rationalizing along similar lines the non-use of the article(s) in such super-compounds as *from coast to coast*, *man-to-man*, *man and woman*, *upstairs* (<=*up the stairs*), *to go to school*, etc.

While we are on the subject of the articles in English, it may be pointed out that osmotic pressure apparently has had a pivotal role to play in the derivation of *the* and *a(n)* from the deictic *that* and the numerical *one* respectively. Always bound to the head nouns they are proclitics to, these words as articles (i.e. *the* and *a(n)*) have had to share

therewith the mass/space of the noun phrases in question as their peripheral constituents. On the other hand, the same words as nominals (i.e. *that* and *one*) have been used as free forms, either of which may occur by itself as the sole occupant of an entire noun-phrase slot. This difference should help explain why *that* and *one* have come to outweigh *the* and *a(n)* respectively in present-day English in both meaning and form, despite their common historical roots.

Incidentally, we can provide an osmosis-based account for the complementary distribution of *one* and *a(n)*, as illustrated in the sentence below.

- (184) *An army* of stags led by a lion would be more formidable than *one* of lions led by a stag.

Note that *one* in the above proverb, but not *an*, is big enough to occupy an entire noun-phrase slot, comfortably accommodating the trace of both the determiner *an* and the head noun *army*. On the one hand, *an* here, but not *one*, is small enough to fit snugly into the peripheral determiner slot to form in conjunction with the head noun *army* a noun phrase that is on a formal par with *one*. Thus *an* and *one* are so distributed here as to help the two noun phrases in question attain a formal equilibrium, arguably under osmotic pressure.

Needless to say, an exactly identical account is applicable to the *a/one* alternation exemplified by the following paraphrase pair.

- (185) a. I'll be back in *a day* or two.
b. I'll be back in *one* or two days.

Rationalizable in essentially the same way as the derivation of *the* and *a(n)* from *that* and *one* respectively is the historical divergence of one and the same negative word into *nought* and *not*. Let us examine the following set of sentences with particular reference to the italicized portions.

- (186) a. *Nought* is never in danger.
b. John is *not* interested in our program.
c. *Not* a single neighbor showed up.

We can see here that *nought* is a free form that occurs all by itself as the sole occupant of a noun-phrase slot, whereas *not* is a clitic that is peripherally bound to the head of the phrase in which it occurs. Thus *nought* takes up an entire phrasal slot, while *not* takes up only a limited portion thereof. This is evidently why *nought* has come to outweigh *not*, (not just semantically but also) formally and ortho-phonologically.

An essentially identical analysis applies to the historical divergence of one and the same Old English word *thurh* into the disyllabic adjective *thorough* and the monosyllabic preposition *through*. The adjective form here is a free form that is capable occurring by itself as the sole constituent of an adjective phrase with plenty of mass/space assigned to it. As a preposition, on the other hand, it is a proclitic that is always peripherally bound to its complement so that it ends up with much more limited mass/space assigned to it. Thus the Old English *thurh* has osmotically evolved into the longer, adjectival form *thorough* and the shorter, prepositional form *through*.

Our discussion here may also provide a new perspective on alternations between affixes and prepositions/particles such as those illustrated in the paraphrase pairs below.

- (187) a. We were *awaiting* the arrival of our equipment.
 b. We were *waiting for* the arrival of our equipment.
- (188) a. The world of books is something *afar*.
 b. The world of books is something *far away*.

The prefix *a-* and the preposition *for* compensate for each other in the first pair here, as do the prefix *a-* and the adverbial particle *away* in the second pair. Osmotic pressure arguably triggers this mutual compensation, which is often observed between affixes and their preposition/particle equivalents.

Analyzable along comparable lines is the alternation between *arise* and *rise from the dead*, as illustrated by the following two sentences taken from *The New American Bible*.

- (189) a. When they *arise*, whose wife will she be? (*Mark 12:23*)
 b. When they *rise from the dead*, they neither marry nor are given in marriage, but they are like angels in heaven. (*Mark 12:25*)

It is interesting to note that the prefix *a-* in *arise* here is arguably a compensatory trace, as it were, for the adverbial *from the dead* in *rise from the dead*.

The alternation between *to awake* and *to wake up*, as illustrated by the following paraphrase pair, may be seen in practically the same light.

- (190) a. I *awoke*, feeling that someone was nearby.
 b. I *woke up*, feeling that someone was nearby.

Incidentally, *up* here, though a word orthographically, may be regarded as something of an affix, i.e. a quasi-affix, since it is equivalent to the prefix *a-*. Besides, *to wake up* is functionally one word so that *up* here may not really be an independent word but only (a peripheral) part thereof. A fairly detailed discussion of the particle *up*, including its affixal nature, is presented in Park (2001).

In fact, the particle *up* seems to frequently figure in osmosis-conditioned alternations of the sort under discussion, as can be attested to by paraphrase pairs such as the following.

- (191) a. Try and avoid getting *entangled* in their vendetta.
 b. Try and avoid getting *tangled up* in their vendetta.
- (192) a. She's decided to *confess*.
 b. She's decided to *fess up*.
- (193) a. Our EFL program needs to *be revamped*.
 b. Our EFL program needs to *be vamped up*.

The particle *up* here evidently fills in for the prefix *en-* in the first paraphrase pair, for the prefix *con-* in the second, and for the prefix *re-* in the third, apparently under osmotic pressure. Also, being on a formal par with the prefixes *en-*, *con-* and *re-*, the particle *up* here may very well be affixal in nature.

It is interesting that the complementary distribution between the prefix *re-* and the particle *up*, as exemplified by (193) above, is apparently fairly productive. In support of this contention, we may cite examples such as the following.

- (194) a. A workout leaves me feeling *recharged*.
 b. A workout leaves me feeling *charged up*.
- (195) a. Airplane factories were able to *retool* quickly for expanded production.
 b. Airplane factories were able to *tool up* quickly for expanded production.
- (196) a. The movie is a *retouched* version of the rags-to-riches story of his life.
 b. The movie is a *touched-up* version of the rags-to-riches story of his life.

Suffixes also may serve to help equivalent cognate phrases to achieve formal equilibrium with each other. For example, the quasi-plural suffix *-s* often serves as a compensatory, osmotic fill-in for what gets deleted from a longer expression, as can be seen from paraphrase pairs such as the following.

- (197) a. Hollands <= Holland *gin*
 b. gramps <=grampa (<=grandpa)
 c. sweats <= sweatsuit
 d. the Fabs <= the *Fabulous Four* ("The Beatles")
 e. vibes <= vibraphone
 f. Beds <= *Bedfordshire*

The quasi-plural suffix *-(e)s* here arguably compensates for *gin* in (197a), *-a* in (197b), *-suit* in (197c), for *-ulous Four* in (197d), for *-raphone* in (197e), and *-fordshire* in (197f).

Note that in (198) below the suffix *-s* is most likely a quasi-plural suffix in *pipes* and something of a plural suffix in *bagpipe(s)* and *panpipe(s)*, where either *pipes* or either of its source forms refers to a (single) wind instrument.

- (198) pipes <= bagpipe(s) / panpipe(s)

It goes without saying that the obligatory suffix *-s* in *pipes* is apparently an osmotic fill-in for *bag-(s)* or *pan-(s)* as the case may be. The optional

suffix *-s* in *bagpipe(s)* or *panpipe(s)*, on the other hand, is very much like a genuine plural suffix in that it is due to the musical instrument in question comprising more than one pipe.

Our discussion here seems to provide a rationalization for the quasi-plural suffix *-s* that figures in the the first member of each paraphrase pair below.

- (199) a. *Barbra (Streisend)* sang the U.S. national anthem at the ceremony.
 b. *Babs* sang the U.S. national anthem at the ceremony.
- (200) a. *(Andre) Agassi* was once married to Brook Shields.
 b. *Ags* was once married to Brook Shields.
- (201) a. *(Prince) William (of Wales)* is on the shy side like his mother.
 b. *Wills* is on the shy side like his mother.
- (202) a. *Gugliotta* is out for the remainder of the season.
 b. *Guggs* is out for the remainder of the season.

The name-final suffix *-s* in *Babs*, *Ags*, *Wills* and *Guggs* here is an osmotic fill-in for *-ra* (*Streisend*), *-assi*, *(Prince) -iam (of Wales)* and *-liotta* respectively. Note incidentally that the suffix *-s* is supplemented by the extra letter *g* in *Guggs* for *Gugliotta*.

Similarly construable is the (quasi-)plural suffix *-s* that figures in the italicized portions of the following paraphrase pairs

- (203) a. My favorite MLS team is *the New England Revolution*.
 b. My favorite MLS team is *the New England Revs*.
- (204) a. *The Colorado Avalanche* swept the Detroit Red Wings in the playoffs.
 b. *The Colorado Avs* swept the Detroit Red Wings in the playoffs.

It may be suggested in this connection that *the Montreal Expos*, the name of a Montreal- based Major League baseball team, arguably derives

from *the Montreal Exposition* under identical osmotic pressure. Parenthetically, the *the* - *s* tandem in the abbreviated forms in question here also is evidently intended to be redolent of the set-denoting *the* - *s* tandem in *the Pittsburgh Steelers*, *the UCLA Bruins*, etc.

The osmosis-triggered suffix *-s* in question may be what figures in such quasi-plural nouns as *news* and *remains*, as is suggested by the following paraphrase pairs.

- (205) a. news <= what is new
b. remains <= what remains

We are suggesting that the suffix *-s* here serves as an osmotic fill-in for something like *what is* in (205a) and *what -s* in (205b).

The suffix *-s* in *Thanks*, *Best wishes* and *Congratulations* also may be similarly triggered by osmotic pressure, as is suggested by the following paraphrase pairs.

- (206) a. Thank you. / I thank you.
b. Thanks.

- (207) a. I wish you the best.
b. Best wishes.

- (208) a. I congratulate you.
b. Congratulations.

The speculation here is that the quasi-plural suffix *-s* is arguably a deletion-compensating fill-in for something like (*I*) - *you*. Note here that the suffix *-s* is supplemented by the noun-forming suffix *-ion* in compensating for *I - you* in (208). In fact, the suffix *-s* is thus supplemented by *-ion* and other noun-forming suffixes in numerous other instances, e.g. *Felicitations* (<=I felicitate you), *Greetings* (<=I greet you), *surroundings* (<=what surrounds (us)) and *environ(ment)s* (<=what environs (us)).

Sometimes the quasi-plural suffix *-s* seems to work in tandem with the definite article, as can be seen from the paraphrase pairs below.

- (209) a. *Gonzaga University* was the favorite in the tournament.
 b. *The Zags* were the favorite in the tournament.
- (210) a. Dorothy Hamil was once the queen of *figure skating*.
 b. Dorothy Hamil was once the queen of *the figs*.

Incidentally, the use of the *the* - *s* tandem in (209b) here as well as in (210b), especially the former, may be motivated at least in part by the meaning of set membership that it implies. The *the* - *s* tandem in *the Zags* is apparently patterned on the *the* - *s* tandem that figures in names of athletic teams such as *the LA Lakers*, *the New York Yankees* and *the Chicago Bears*.

The *the* - *s* tandem as an osmotic fill-in may not always be associated with the meaning of set membership, as can be seen from paraphrase pairs such as the following.

- (211) a. We were at *a movie theater/house* on Saturday night.
 b. We were at *the movies* on Saturday night.
- (212) a. We were at *a picture theater/house* on Saturday night.
 b. We were at *the pictures* on Saturday night.

The *the* - *s* tandem in the second member of each pair above is arguably an osmotic fill-in for *a* - *theater/house* in the first member of the same pair.

We are now in a position to provide a highly plausible account for the origin of the quasi-plural suffix *-s* used in the second member of the following paraphrase pair.

- (213) a. She's *on top* as far as I'm concerned.
 b. She's *(the) tops* as far as I'm concerned.

Assuming that *(the) tops* here derives from *on top*, which is quite plausible, we may argue that *(the) -s* is compensatory for the preposition *on*, which gets deleted in the derivation of *(the) tops* from *on top*. The underlying source for *(the) tops* here may alternatively be posited as *at the top*, in which case *(the) -s* is a fill-in for *at the*. Indeed, the suffix *-s* often apparently compensates for other prepositions than *on* as well, e.g.

for *at* as in *He works nights* for *He works at night*. It is tempting to speculate at this point that the suffix *-s* in the adverb *besides* is an osmotic fill-in for the complement of the preposition *beside*.

Our discussion here may also throw light on the differential dispensability of the experiencer phrase displayed by the two members of either paraphrase pair below.

- (214) a. The crime rate here *is frightening* (to us).
 b. The crime rate here *frightens* us.
- (215) a. The display *was impressive* (to us).
 b. The display *impressed* us.

On the one hand, the verb phrases (*is*) *frightening* and (*was*) *impressive*, which begin with *to be* and terminate with their respective derivational suffixes *-ing* and *-ive*, may already be sufficiently saturated spatially to normally render the experiencer phrase *to us* almost an “unwelcome” appendix. On the other hand, the verb phrases *frightens* and *impressed*, which terminate with just their respective inflectional suffixes *-s* and *-ed*, may be spatially so sparse that the experiencer phrase *us* may normally serve the useful purpose of bulking up the verb phrase to a desired level of formal mass/space. Arguably osmosis-triggered, this difference seems to help explain why the experiencer phrase is less obligatory in the first member of either pair above than in the second.

Our discussion here also may be able to provide a vantage point from which to explain why in the following pair of paraphrase sentences *face* requires reinforcement by *of it*, which *surface* disallows.

- (216) a. *On the face of it*, it sounds like a good idea.
 b. *On the surface*, it sounds like a good idea.

Note that *face* is *surface* minus the prefix *sur-* so that the former is considerably lighter/shorter than the latter formally (and perhaps semantically as well). Thus in the sentence-initial adverbial in question here, *surface* may already be spatially so heavy relative to *face* that the former can ill afford reinforcement by *of it*, which the latter cannot afford to do without. That is, if the sentence-initial adverbial with *surface* here is to be on a formal par with its equivalent with *face*. Thus it may

very well be that *of it* and *sur-* osmotically counterbalance each other here to put the two adverbials on a mutual formal par.

Our discussion here also may shine some interesting light on the alternation between *by far* and *far and away*, as illustrated by the following paraphrase pair.

- (217) a. Longman is *by far* the best-known dictionary publisher.
 b. Longman is *far and away* the best-known dictionary publisher.

It would appear here that *by* (in *by far*) and *and away* (in *far and away*) are mutually exclusive, osmotically canceling each other out, so that *by far* and *far and away* are arguably on a formal par.

5. Semantic Osmosis

The structure of meaning also appears to be subject to osmotic pressure such that the semantic mass/space of a linguistic expression tends to vary inversely with that of the material with which it is in construction. Let's begin our discussion here with a look at the sentence pairs below, focusing on the underlined verbs.

- (218) a. What did you ask *him*?
 b. What did you ask *him to do*?
- (219) a. I found *the book*.
 b. I found *the book to be readable*.
- (220) a. She kept *the coins*.
 b. She kept *the coins in mint condition*.
- (221) a. He took *the lamp*.
 b. He took *the lamp to the upstairs bedroom*.
- (222) a. He made *my torch*.
 b. He made *my torch longer and better*.

- (223) a. He pronounced the verdict.
 b. He pronounced the verdict appropriate.
- (224) a. Who told him?
 b. Who told him to go?

Note here that the semantic mass/space of the main verb varies from one member of each pair to the other, roughly in inverse proportion to the semantic mass/space of its complement, i.e. the remainder of the verb phrase. Thus one and the same verb *to ask*, for one, is apparently (more concrete and thus) heavier in meaning in (218a) than in (218b) such that *to ask him* in the former is arguably equivalent in semantic mass/space to *to ask him to do* in the latter.

An exactly identical phenomenon seems to be observable in the italicized portions of the sentence pairs below.

- (225) a. I'm going.
 b. I'm going nuts.
- (226) a. He's going to come (and tell you everything).
 b. He's going to come clean (and tell you everything).
- (227) a. The grass is growing.
 b. The grass is growing greener.
- (228) a. He appeared at the conference.
 b. He appeared to be busy at the conference.
- (229) a. He spoke at the meeting.
 b. He spoke English at the meeting.

It is interesting that our discussion here apparently helps explain why one and the same verb is semantically heavier as a main verb than as an auxiliary, as can be seen from a consideration of the following pairs of sentences.

- (230) a. We have lots of books here.
 b. We have bought lots of books here.

- (231) a. We are in a nice hotel.
 b. We are *accommodated* in a nice hotel.
- (232) a. He wants to will all his wealth to his wife.
 b. He will *leave* all his wealth to his wife.

Have, *are* and *will* as main verbs here monopolize the semantic mass/space of the verb phrase proper, while the same verbs as auxiliaries share in a peripheral capacity the same semantic mass/space with their respective main verbs. (Incidentally, the verb phrase proper here refers to the verb phrase up to the main verb and thus may comprise either an auxiliary (string) or a main verb or both.) Thus the verb in question here has plenty of semantic mass/space to account for in the first member of each pair here, as opposed to much more limited semantic mass/space in the second member of the same pair. This is arguably a major reason the verb in question ends up being heavier in meaning in the first member of each pair here than in the second. As we shall see toward the end of this section, cumulative osmotic pressure also plays a role in rendering one and the same verb here semantically (and thus prosodically) more lightweight as an auxiliary verb than as a main verb.

It is in order here to point out that even auxiliary verbs sometimes carry as much semantic weight as, or even more semantic weight than, main verbs do. We have already suggested in Section 1 that this is the case, for example, when auxiliaries are clause-final. A clause-final auxiliary is heavy in both meaning and (phonological) form because as the sole constituent of a whole verb phrase it osmotically absorbs the extra semantic weight of its underlying complement over and above its own semantic weight

Our discussion here apparently makes it possible to explain in a principled way why the first token of *is* in the sentence below is heavier in meaning than the second.

- (233) Whatever *is*, *is* right.

The verb phrase of the subject clause here comprises only the verb *is*, while the verb phrase of the sentence itself comprises the verb *is* plus its complement *right*. Thus the first token of *is* here takes up the semantic mass/space of a whole verb phrase, whereas the second token has to

share the same mass/space with its complement. This, in our terms, may be the main reason why the first token of *is* is heavier than the second token of the same here semantically (and thus prosodically as well).

An exactly identical account is applicable to why the verb *to be* in italics is heavier in meaning in the first member of the following pair than in the second.

- (234) a. *To be or not to be*, that is the question.
 b. It is politically correct these days *to be environment-friendly*.

We can rationalize along essentially identical lines why the first token of the verb *to do* is heavier in meaning than the second in (235) below.

- (235) a. Those who can, *do*; those who can't, teach.
 b. They are all very eager to *do their part*.

Our discussion also seems to make possible a principled explanation for the difference in semantic mass/space between the two tokens of *have*, as used in the following paraphrase pair.

- (236) a. The tax system allegedly favors *those who have wealth*.
 b. The tax system allegedly favors *the haves*.

Assuming that *the haves* is derived from *those who have wealth* here, we may argue that the meaning of *the haves* osmotically contains the meaning of not just *have* but also *those who - wealth*. Thus of the two tokens of *have* here, the second is heavier in meaning than the first. Evidently reflective of this difference in semantic mass/space, *have* here is in construction with a shorter stretch of material and is thus osmotically less concessive of meaning thereto in *the haves* than in *those who have wealth*. Incidentally, we occasionally come across *the Haves*, in which the uppercase initial *H* is arguably intended as an orthographic reflex of the semantic promotion of *have* in the verb-to-noun derivation here.

From a purely formal perspective, the *the - -s* tandem in *the haves* appears to be a fill-in for *those who - wealth* in *those who have wealth*. However, it may be more accurate to view the *the - -s* tandem here as a fill-in for just *those who* with *wealth* fusing semantically with *have* in

the haves. This semantic incorporation of *wealth* into *have* should help account for the greater semantic (and thus prosodic) weight of *have* in *the haves* than in *those who have wealth*.

The tendency toward an osmosis-conditioned semantic equilibrium of basically the same nature may also be observed in sentence pairs such as the following.

- (237) a. You must be crazy.
 b. You must be crazy about music.
- (238) a. She's good.
 b. She's good at tennis.
- (239) a. He was full.
 b. He was full of bigotry.

The semantic mass/space of the adjective in each pair here is apparently in inverse proportion to that of its complement.

Explainable along essentially identical lines is the difference in semantic mass/space between the two tokens of the underlined word in each sentence pair below.

- (240) a. It's on *the* bottom of the drawer. (bottom: lowest part)
 b. It's in *the* bottom drawer. (bottom: lowest)
- (241) a. It's near *the* top of the shelf. (top: highest part)
 b. It's on *the* top shelf. (top: highest)
- (242) a. She was in *the* middle of a room. (middle: central part)
 b. She was in *the* middle room. (middle: central)

The parenthesized glosses for the two tokens of one and the same underlined word in each pair here clearly show that the word in question is semantically a little heavier in the first member of the pair than in the second. The word in question here is the head of a noun phrase in the first member of each pair and thus may take up much of the (semantic) mass/space assigned to the phrase. The same word is an attributive modifier for the head of a noun phrase in the second member

of the same pair and thus gets a more limited share of the (semantic) mass/space assigned to the phrase. This apparently underlies the observed difference in semantic mass/space between the two tokens of the underlined word in each pair here.

An exactly identical interpretation applies to the difference in semantic mass/space between the two tokens of the underlined word in each sentence pair below.

- (243) a. The couch was in *the front* of the room.
 b. The couch was in *the front room*.
- (244) a. They keep the door at *the back* locked.
 b. They keep the *back door* locked.
- (245) a. This slide shows the building, as viewed from *the side*.
 b. This slide shows *a side view* of the building.
- (246) a. Use the door in the *rear*.
 b. Use the *rear door*.

Osmotic pressure also apparently underlies the variation in semantic mass/space between the two cognate tokens of the underlined word in each of the following pairs.

- (247) a. *Pretty!*
 b. *Pretty good!*
- (248) a. *Awful!*
 b. *Awful cold!*
- (249) a. It's *real*.
 b. It's *real good*.
- (250) a. The weather is *nice and balmy*.
 b. The weather is *nice*.

Note that *nice and* in (250a) is intended as an idiomatic intensifier and that the variation in semantic mass/space involving *nice* in (250) would

remain the same even if *nice* were replaced with *good*.

Let us now examine the following set of sentences with particular reference to the semantics of the word *that*, which is used as a demonstrative pronoun in the first sentence and as a connective in the remaining two sentences.

- (251) a. *That* is a nice car.
 b. I know *that* he drives a nice car.
 c. I know the nice car *that* he drives.

Other things being equal, the demonstrative *that* is heavier in meaning (and thus in pronunciation as well) than the connective *that*. This is apparently mainly because the demonstrative *that* has to itself the semantic slot for an entire (noun) phrase here, while the connective *that* has to share the semantic slot for a (connected/"conjoined") phrase with the two stretches of material it serves to join together.

We can offer an essentially identical argument as to why the *WH*-word in each sentence pair below is heavier in meaning (and thus in pronunciation as well) in the first member than in the second.

- (252) a. *Who* do you work for?
 b. There are those *who* eat out for a special occasion.

- (253) a. *What* do you want?
 b. She is in *what* the doctors call a hypochondriac.

- (254) a. *When* will he arrive?
 b. I'll leave *when* he arrives.

Explainable along similar lines is the difference in semantic mass/space between the two tokens of the word in italics in each sentence pair below.

- (255) a. That's a big *if*.
 b. I'd go *if* I were you.

- (256) a. No *but*s, you're going to school today.
 b. I'd like to go, *but* I'm awfully busy today.

The de-conjunctival nouns here may occur as heads of (noun) phrases (sometimes all by themselves), while the conjunctions from which they stem are bound to the two stretches of material that they serve to tie together. We may say that the de-conjunctival nouns here osmotically absorb the mass/space not just for the conjunctions from which they derive but also for (the traces of) whatever these underlying conjunctions serve to bind together. This is apparently why the de-conjunctival nouns here are normally heavier than their respective conjunctival sources semantically (and thus prosodically as well). This difference in semantic mass/space is evidently reflected in surface form by the de-conjunctival nouns here being in construction with shorter stretches of material and thus osmotically less concessive of meaning than are their respective conjunctival cognates.

We can explain along essentially identical lines why one and the same word is normally less heavy semantically (and thus prosodically as well) as a preposition than as a particle, as can be seen from pairs of paraphrase sentences such as the following.

- (257) a. He's *in* the house.
b. He's *in*.

- (258) a. Keep *off* the grass.
b. Keep *off*.

Note that the particle in the second member of each pair here stands for the prepositional phrase in the first member of the same pair and thus osmotically bears the semantic mass/space of not just the preposition but also (the trace for) its complement. This is evidently why the particle in the second member of each pair here is heavier semantically (and thus prosodically as well) than its prepositional counterpart in the first member.

We also can cite paraphrase pairs such as the following in further illustration of osmotically-conditioned semantic variance of the sort under discussion.

- (259) a. Did you read *some/any/both* books?
b. Did you read *some/any/both*?

- (260) a. We will use *either/neither* book.
 b. We will use *either/neither*.

In each pair here, the one-word direct object in the second member derives from its two-word counterpart in the first. Note that the two equivalent direct objects here have one word in common and that this shared word in the second member of each pair osmotically absorbs the mass/space of not just its cognate in the first member of the same pair but also (the trace for) the head noun that it predetermines. This should help explain why the second token of the shared word in question here is heavier than the first thereof semantically (and thus prosodically as well).

It is interesting to note here that the attributive adjective *few* is heavier semantically (and thus prosodically as well) when it occurs by itself than when it is predetermined by the indefinite article *a*. With this in mind, let us compare the two italicized noun phrases in the following pair of sentences.

- (261) a. *Few* people came to the concert.
 b. A *few* people came to the concert.

Note that *few* gets all the semantic mass/space assigned to the attributive modifier slot in the first member of the pair here, while it has to share this mass/space with the indefinite article *a* in the second member of the same pair. As a result, *few* is heavier in meaning (and thus in prosody as well) in the first member of the above pair than in the second. Our discussion here assumes that negativity outweighs positivity in terms of semantic content.

Needless to say, we can explain in exactly the same way why the attributive adjective *little* is heavier semantically and (thus prosodically as well) in the first member of the sentence pair below than in the second.

- (262) a. *Little* attention was paid to the matter.
 b. A *little* attention was paid to the matter.

Given examples like (261) and (262), the indefinite article sometimes appears to play the role of semantic-mass/space diluter. It is, in fact,

apparently as much an adjuster of semantic mass/space as it is one of formal-mass/space. Recall that we have already discussed the (in)definite article as adjuster of formal mass/space in the preceding section.

Our current discussion seems to cast light on why *no* is heavier in meaning than *not* in the context of a paraphrase pair such as the one below.

- (263) a. The food was no good.
 b. The food was not good at all. / The food was not any good.

No good being equivalent in meaning to *not good at all* or *not any good*, *no* here osmotically accommodates the meaning of *at all* or *any* over and above the negative meaning of *not*, so to speak, so that *no* is apparently heavier in meaning than *not* in the paraphrase pair above. Note here that *no* is in construction with a shorter stretch of material than *not* is, which is apparently a surface reflex of the former making less osmotic concession of meaning than the latter. Incidentally, given our discussion here, it is quite clear that “*The food was no good*” does not mean the same thing as “*The food was not good*.”

Our current discussion seems to provide a fresh angle from which to view the semantic alternation between the indefinite article *a* and the numeral *one* in a minimal pair such as the following.

- (264) a. We needed a maid to clean the room.
 b. We needed one maid to clean the room.

Note that the information focus here is on *maid* in *a maid* and on *one* in *one maid* and that the focused word here osmotically gains semantic/informational weight at the cost of the “unfocused” word in the same phrase. Note also that the focused *one* is heavier than the unfocused *a*, the former making less osmotic concession of meaning and form (than does the latter) to the head noun *maid* that it predetermines here. This range of osmotic variations apparently results in the two noun phrases here achieving a balance of semantic/informational mass/space.

An identical analysis may be imposed on the alternation between *the* and *that*, as illustrated in a minimal pair such as the following.

- (265) a. We needed *the maid* to clean the room.
 b. We needed *that maid* to clean the room.

The information focus is on *maid* in *the maid* here and on *that* in *that maid*, so that *maid* is heavier in information/meaning content in the former than in the latter. Since the focused demonstrative *that* is heavier in meaning (and form) than the unfocused article *the*, the variation here evidently helps the two noun phrases in question to achieve a mutual osmotic parity of meaning.

Osmotic pressure might also be at work in derivation-induced dilution of meaning, as is apparently illustrated by examples such as the following.

- (266) a. Most people prefer *red* roses.
 b. There was a bonfire blazing *redly*.

It appears that *red* is greater in semantic mass/space in the first sentence here than in the second in that it retains more of its original color meaning in the former than in the latter. If this is correct, then it arguably has to do with *red* taking up the entire semantic slot of a word in the first sentence, where it is an adjective, as opposed to just a part of one in the second sentence, where it is merely the stem of an *-ly* adverb. *Red* is an independent word in (266a) and thus has no (intra-lexical) osmotic concession of meaning (and pronunciation) to make, whereas it is part of the word *redly* in (266b) and thus has some (intra-lexical) semantic (and phonological) concession to make to *-ly*.

We may point out in this connection that historically the suffix *-ly* has made far more concession of meaning (and thus pronunciation as well) to the stem *red* here than the other way around. Like all suffixes, in fact, *-ly* has had so much concession to make that it is hardly recognizable today as having anything to do with *like*, which is its etymological ancestor. We will argue later on that much of this concession on the part of an affix, such as the suffix *-ly* here, is cumulative.

This osmotically-triggered concession that the *-ly* suffix and its stem make to each other, be the resulting word adverbial or adjectival, appears to be fairly productive. This is apparently attested to by an abundance of such pairs of cognate words as *near/nearly*, *short/shortly*, *bare/barely*, *hard/hardly*, *low/lowly*, *live/lively*, *lone/lonely*, *clean/cleanly*, and

sick/sickly.

A similar phenomenon of semantic dilution is observable in the noun-to-adjective derivations illustrated by the italicized cognate pairs in examples such as those that follow.

- (267) a. a *gold* watch : a *golden* opportunity
 b. an *ash* tray : an *ashen* face
 c. a *brass* bed : a *brazen* hussy

The adjectival suffix *-en* apparently serves to dilute the original meaning of the stem noun in each of the derivations involved here. Note in this connection that the stem of a word generally seems to get semantically diluted in proportion to the derivational complexity of the word in which it occurs and that this dilution is arguably triggered by osmotic pressure.

This correlation between derivational complexity and semantic dilution seems to manifest itself on the supra-lexical level as well, as can be seen from an examination of the following active/passive pairs of sentences.

- (268) a. They *made* him *leave* the country.
 b. He *was made to leave* the country.

- (269) a. They *saw* her *sneak* upstairs.
 b. She *was seen to sneak* upstairs.

It appears that the active versions here (i.e. *make leave* and *saw sneak*) are expressive of more immediate causation and perception respectively and thus arguably heavier in meaning than are their respective passive versions (i.e. *was made to leave* and *was seen to sneak*). If this is correct, then we can say that the active/passive pairs above are supportive of our claim about the correlation between derivational complexity and semantic dilution. We are assuming here that passives are derived from their active counterparts, i.e. the former are derivationally more complex than the latter.

Note in this connection that the main verb in each pair above apparently shares the verb phrase slot with a shorter stretch of material in the active voice than in the passive. This may be reflective of the main verb making less osmotic concession of causative/perceptual meaning in the active voice than in the passive. Incidentally, the

infinitive-marking *to* (in 268b and 269b) here may very well be a surface reflex of the underlying barrier that serves to keep down the immediacy of causation and perception.

Parenthetically, interpretable along basically identical lines is the difference in perceptual immediacy denoted by the two verb phrases below.

- (270) a. He *was found dead*
 b. He *was found to be dead*.

The verb *found* is in construction with a shorter stretch of material in the first sentence here than in the second. This is arguably a surface reflex of *found* making less osmotic concession of its perceptual force and thus being expressive of greater perceptual immediacy in the first sentence above than in the second.

We may advert here to *to look* denoting more immediate perception than *to seem*, which may be why *to be* is acceptable as complementizer for *to seem*, but not normally for *to look*. Let us compare the following two sentences, focusing upon the internal structure of their verb phrases.

- (271) a. He *looks (??to be) nice*.
 b. He *seems (to be) nice*.

We may argue here that *looks* is expressive of greater perceptual immediacy and, in that sense, semantically heavier than is *seems*, so that *looks* cannot afford quite as much osmotic room for a complementizer as *seems* can. This is arguably why *to be* may be used as complementizer for *seems*, but not normally for *looks*.

The adhesiveness/ cohesiveness with which the constituents of an expression are bound to each other seems to play a role in determining the force with which osmotic pressure is exerted. Let us consider the following pair of sentences, focusing attention on the two noun phrases in italics.

- (272) a. I saw him in *a greenhouse*.
 b. I saw him in *a green house*.

We may note that *greenhouse* is an (idiomatic) lexical combination and

green house a (non-idiomatic) phrasal combination, so that *green* and *house* are more closely attached to each other in the former than in the latter. Thus the two immediate constituents here arguably make more osmotic concession of meaning to each other in *greenhouse* than in *green house*, which is apparently why both constituents are semantically weaker in the former than in the latter.

Note that *a greenhouse* is not necessarily either “*green*” nor is it necessarily “*a house*,” while *a green house* is in fact “*a house that is green*.” Thus *green* and *house* may be said to be more diluted and thus lighter in meaning in the first sentence above than in the second, i.e. in that they are further removed from their original meanings in the former than in the latter. If we may assume that *greenhouse* is derived from *green house*, which does not seem to be too implausible, the semantic dilution/weakening observable here is arguably a function of the former being derivationally more complex than the latter.

We may observe in passing here that the two constituents in question also make more phonological concession to each other in *greenhouse* than in *green house*, which is reflected in them both normally getting phonologically weaker in the former than in the latter. Note, for example, that the the two nuclei /i:/ and /au/ here are normally weaker phonetically in *greenhouse* than in *green house*. Of relevance to the point being made here seems to be the frequent downgrading of the first-syllable nucleus in *Greenwich* from /i:/ to /e/ (and of the *Greenwich*-final affricate from voiceless to voiced), especially in *Greenwich (Mean) Time*.

Such differentially exerted osmotic pressure as is observed in *greenhouse* vs. *green house* seems to apply *mutatis mutandis*, whenever two words may occur in either an (idiomatic) lexical combination or a (non-idiomatic) phrasal one. With this in mind, let us consider the two tokens of *in time* illustrated in the sentence pair below.

- (273) a. They arrived just *in time* for the meeting.
 b. They were separated *in time* by almost two centuries.

The expression *in time* is an (idiomatic) lexical combination in the first sentence here and a (non-idiomatic) phrasal combination in the second. Thus *in* and *time* are more closely bound to each other in the first sentence than in the second so that the two constituents apparently

make more osmotic concession of meaning to each other in the former than in the latter. Note that the original meanings of both *in* and *time* become less distinct and thus less heavy in the (idiomatic) lexical combination than in the (non-idiomatic) phrasal combination.

An exactly identical analysis applies to the two constituents of *on earth* making more osmotic concession of meaning to each other and thus diluting each other's meaning more in the first sentence below than in the second.

- (274) a. Why *on earth* would anyone do a thing like that?
 b. The circus used to be referred to as the greatest show *on earth*.

Let us now compare *reserve* and *re-serve*, as used in the following sentences, focusing on the two tokens of the prefix *re-* with reference to the adhesiveness/cohesiveness with which they are bound to the stem *-serve*.

- (275) a. I'd like to reserve a table for two.
 b. We had to re-serve refreshments.

The etymologically identical prefix *re-* here is more tightly affixed to the stem *-serve* in *reserve* than in *re-serve*, which may result in the affix and stem making more osmotic concession of meaning to each other in the former than in the latter. This should help explain why both the prefix and the stem are much weaker semantically in *reserve* than *re-serve*. In fact, the prefix and the stem make so much semantic concession to each other in *reserve* that they end up practically merging into each other with the result that their separate semantic identities, if any, seem to be hardly discernible. It may very well be that *reserve* is an "idiomatic" combination and *re-serve* a "non-idiomatic" combination, so to speak.

Not coincidentally, the prefix *re-* and the stem *-serve* here make a distinctly greater osmotic concession of phonological mass/space to each other in *reserve* than in *re-serve*. This is apparently why the stem-initial consonant /s/ gets voiced (and thus weakened) to /z/ in *reserve*, but not in *re-serve*, and why the prefix nucleus /i:/ gets weakened to /i/ in the former, but not in the latter.

We can explain along identical lines that the so-called agent suffix *-er*

and the stem are bound more closely to each other and thus make more mutual concession of meaning, in the first member of either cognate pair below than in the second member of the same pair.

- (276) a. *prayer* (“a devout petition to God”)
 b. *prayer* (“a person who prays”)

- (277) a. *drawer* (“a sliding compartment in a piece of furniture”)
 b. *drawer* (“a person who draws”)

Note here that the original meanings of both the stem and the suffix are apparently more diluted in the first member of either pair here than in the second. Note also that there is a phonological consequence of the semantic concession that the suffix *-er* and the stem make to each other in that the two merge into a monosyllable in the first member of either pair above, but not in the second.

Osmotic pressure on the semantic content of a particular element is exerted not just by what it is in immediate physical construction with but also by what it can be in potential construction with. Let us examine the following data with reference to the diversity of what the ‘light(weight)’ verbs in italics may cumulatively enter into construction with.

- (278) a. to *do* the dishes / to *do* the laundry / to *do* the sights / to *do* time / to *do* lunch
 b. to *take* a walk / to *take* a rest / to *take* a break / to *take* a look / to *take* note
 c. to *get* up / to *get* down / to *get* in / to *get* out / to *get* on / to *get* off
 d. to *turn* on / to *turn* off / to *turn* up / to *turn* down / to *turn* in / to *turn* away
 e. to *pull* in / to *pull* out / to *pull* away / to *pull* up / to *pull* over / to *pull* through
 f. to *put* aside / to *put* away / to *put* off / to *put* on / to *put* out / to *put* up
 g. to *make* believe / to *make* as if / to *make* good / to *make* time / to *make* water

Each verb here can potentially enter into construction with a great diversity of material so that it arguably ends up making a substantial cumulative concession of meaning thereto, which is apparently the main reason it is normally light of semantic content.

Note that the salient semantic contrast between the members of each verb-phrase set here is signaled by what follows the verb, not by the verb itself, which is common to all the member phrases. The verb here is like an affix and thus light of meaning, whereas its complement is like a stem and thus pregnant with (contrastive) meaning. Parenthetically, this difference in meaning between the verb and its complement underlies the difference in accentuation between the two, helping explain why the verb normally receives much less accentuation than its complement.

Similarly explainable in large part is the typically light meaning/ accentuation of such function words as articles, prepositions, conjunctions and auxiliary verbs. The articles, for example, may enter into construction with a virtually unlimited number of nominals, so that they must make a great cumulative osmotic concession of meaning thereto, which is apparently why they do not carry too much semantic/ phonological weight.

Note in this connection that the semantic weight of a word or, for that matter, any other linguistic element, is in inverse proportion to the frequency with which it occurs. Semantically lightweight words, for example, are high-frequency words belonging to closed classes of words, i.e. function words. These words are used with high frequency because they are capable of occurring in construction with a great diversity of complements. Given these observations, the tendency for high-frequency words to be lightweight in meaning (and thus in pronunciation as well) may be a natural osmotic consequence of their potential to enter into cumulative construction with a great diversity of material.

The relative frequency of an individual linguistic element, be it lexical or otherwise, does indeed seem to translate into its semantic levity in a fairly straightforward manner. Of the two cognate prepositions *of* and *off*, for instance, the former is far more frequent and thus semantically far lighter than the latter. Again, of the two synonymous nouns *thing* and *affair*, the former is far more frequent and thus semantically far lighter than the latter. What is significant here from our point of view is that all this is apparently attributable to the operation of osmotic pressure in the English language.

Our discussion here applies just as well to the sub-lexical level also, as can be seen from a consideration of examples such as the following.

- (279) a. *England / Scotland / Ireland / Finland / Iceland*
 b. *Plymouth / Dartmouth / Portsmouth / Tynemouth*

Note that *-land* and *-mouth* are sub-lexical formatives of relatively high frequency that co-occur with a considerable diversity of material so that they cumulatively make a substantive osmotic concession of semantic (and thus phonological) mass/space thereto. In fact, *-land* and *-mouth* get thus so drained of semantic (and phonological) substance that they are almost like affixes.

Note that *-land* and *-mouth* in (279) is much more tightly affixed to their stems than are the superficially identical *-land* and *-mouth* in (280) below.

- (280) a. *farmland / marshland*
 b. *bigmouth / bad-mouth*

Being much more closely bound to their “stems” in (279) than in (280), *-land* and *-mouth* apparently make much more osmotic concession of meaning (and pronunciation) thereto in the former than in the latter. Thus *-land* and *-mouth* retain much more of their original semantic/phonological values and are thus heavier in meaning/pronunciation in (280) than in (279).

Speaking of sub-lexical examples of relevance to our discussion here, we may note that affixes are more frequent than stems because the former may occur in construction with a greater diversity of the latter than the other way around. Thus affixes make more osmotic concession of meaning (and pronunciation) to stems than the other way around, which is why stems normally outweigh affixes in semantic/phonological mass/space.

Our discussion here also is apparently instrumental in explaining why an utterance is normally (more easily recoverable and thus) more readily intelligible from its consonantal skeleton than from its vocalic skeleton. Let us compare the two skeletons below of one and the same four-word sentence, in which each blank represents a deleted “grapheme.”

- (281) a. i i y a.
 b. Ths s m bg.

It appears that the consonantal skeleton (281b) above provides a much better clue to the intended sentence “*This is my bag*” than does the vocalic skeleton (281a). Note that vowels normally enter into construction with a greater diversity of consonants than the other way around so that vowels make more cumulative osmotic concession of significance/information to consonants than the other way around. This is evidently why consonants are normally more pregnant with significance/information than are vowels.

It is interesting that our discussion of osmotic pressure here should point to the apparent semantic relevance of phonological segments. In fact, we may very well be able to show that the frequency of a phonological segment varies inversely with its semantic/informational contribution to communication. For one, the schwa, by far the most frequent vowel in English, is arguably the least significant English vowel, that is, semantically and informationally. It would be quite interesting to do a comprehensive frequency-based analysis of all the individual sounds of English with reference their differential semantic/informational significance.

It may be speculated at this point that an element also may be osmotically affected by what it is in “referentially paradigmatic” construction with. Take, for example, the two tokens of *Lothario* in the sentence pair below.

- (282) a. He was cast as *Lothario* in the play.
 b. He was cast as *a Lothario* in the play.

Lothario refers to the one and only member of a set of people called “Lothario” in the first sentence, and to one of more than one member of a set of people called “Lothario” in the second sentence. Thus *Lothario* is in “referentially paradigmatic” construction with zero in the first sentence and with non-zero in the second, so to speak. Referring to a whole set in the first sentence above and to only part thereof in the second, *Lothario* is arguably heavier (or less diluted) in referential meaning in the former than in the latter. This difference in semantic mass/space seems to be reflected in surface form here with *Lothario* filling the whole of a noun

phrase slot in the first sentence and only part thereof in the second with the remainder being filled by *a*.

6. Closing Remarks

We have amply demonstrated that osmotic pressure operates quite extensively in the English language in such a way that it serves to help the language achieve and maintain parity-based stability on various levels of its structure. Osmotic pressure, which appears to be conducive to conservation of the mass/space for such elements as are affected, has been shown to be instrumental in rationalizing numerous, often otherwise quite baffling, aspects of the structure of English.

It seems to be in order here to point out that osmotic pressure may be operative in languages other than English as well. For example, osmotic pressure seems to be responsible for the alternation between /hana/ and /han/, the two alternants of the Korean equivalent of the English numeral *one*. We use /hana/ as a head-noun numeral that can stand by itself and /han/ as a proclitic numeral that has to precede a head noun, so that the former gets more mass/space in meaning and form than does the latter. This alternation in Korean is not unlike the *one/a(n)* alternation in English in that both are apparently governed by one and the same osmotic principle.

Far from being specific to English then, osmotic pressure may very well be a linguistic universal. If it so turns out, a study of osmotic pressure as it applies to English will certainly help sharpen our insight into its impact on the structure of English in particular and of human language in general. This seems to call for an extensive cross-linguistic investigation into the phenomenon under discussion here in the years ahead. Such cross-linguistic inquiry may result in a giant step forward in our efforts at a better understanding of human language.

Finally, it should be borne in mind that the present paper represents a work in progress, rather than a completed work. Still in its embryonic stages, the research that has gone into the current paper has barely scratched the surface of the matter under study and is very likely flawed by errors of various sorts, of which we may or may not be aware at this point in time. However, this should not detract from the importance of osmotic pressure as a pervasive force in the structuring of English and

other languages. Rather, it should serve as a motivator for further research into the fascinating world of osmosis as it relates to the arcane inner workings of English as well as other natural languages.

References

- Bronstein, A. J. (1960). *The Pronunciation of American English: An Introduction to Phonetics*. New York: Appleton-Century-Crofts Inc.
- Chafe, W. (1970). *Meaning and the Structure of Language*. Chicago: University of Chicago Press.
- Chomsky, N. and M. Halle. (1968). *The Sound Pattern of English*. New York: Harper and Row.
- Fowler, H. W. (1965). *A Dictionary of Modern English Usage*. Oxford: Oxford University Press.
- Francis, N. (1958). *The Structure of American English*. New York: Ronald Press.
- Hawking, S. (1988). *A Brief History of Time*. New York: Bantam Books.
- Jespersen, O. (1909-1949). *A Modern English Grammar on Historical Principles. Parts I-VII*. Copenhagen: Munkskaard.
- Kreidler, C. W. (1989). *The Pronunciation of English*. Oxford: Basil Blackwell.
- Marchand, H. (1969). *The Categories and Types of Present-Day English Word-Formation*. Munich: Verlag C. H. Beck.
- Park, N.-S. (1977). Structural density as a comprehensibility variable. *Language Research* 13(2), 95-103.
- Park N.-S. (1983). Proximity and reduction. *Language Research* 19(1), 75-102.
- Park, N.-S. (1992). The law of conservation and linguistic explanation. *Language Research* 28(2), 267-334.
- Park, N.-S. (2001). Four faces of *up* as a functional particle. *Language Research* 37(2), 227-265.
- Park, N.-S. (in Preparation). Osmotic depletion in structural condensation.
- Quirk, R., Greenbaum, S., Leech, G., and J. Svartvik. (1985). *A Comprehensive Grammar of the English Language*. London: Longman Group.
- Sinclair, J. (1995). *Collins Cobuild English Dictionary*. London: HarperCollins Publishers.
- Summers, S. (2000). *Longman Advanced American Dictionary*. London:

Longman.

The Good News Bible. (1976). New York: American Bible Society.

The New American Bible. (1991). Michigan: World Publishing, Grand Rapids.

The New Jerusalem Bible. (1990). New York: Double Day.

Nichols, W. R., Editorial Director (2000). *Random House Webster's College Dictionary*. New York: Random House.

Nahm-Sheik Park

Department of English Language and Literature

Seoul National University

San 56-1 Sillim-dong, Gwanak-gu

Seoul 151-742, Korea

E-mail : nspark@snu.ac.kr

Received: Mar. 25, 2002

Revised version received: May. 27, 2002

Accepted: May. 31, 2002