

# Osmotic Depletion in Structural Condensation with Special Reference to Word Formation in English\*

Nahm-Sheik Park  
(Seoul National University)

**Park, Nahm-Sheik. (2004). Osmotic depletion in structural condensation with special reference to word formation in English. *Language Research* 40(2), 313-351.**

As basic building blocks of language, words need to be compact so that they are under rather strict spatial constraints. Thus a supralelexical entity often gets depleted when it morphs into a word, i.e. when it gets (end) osmosed into the cramped space of a word. This seems to be especially the case in the formation of compound words and of derivatives thereof. Building upon Park (1992, 2001, 2002, 2003), the present paper takes a look at this (end)osmotic depletion and related phenomena with reference to English.

**Key words:** English word formation, compound words, depletion, simplification, blending, clipping

## 1. Introduction

For all practical purposes, the word is the smallest functional unit of a language and as such it is assigned only minimal space. Thus a supralelexical entity such as a phrase often gets maximally compressed in one way or another when it metamorphoses into a word. This downsizing is arguably designed to fit the supralelexical entity in question into the cramped space of the resulting word. Building on relevant points made in Park (1992, 2001, 2002, 2003), we will investigate this phenomenon of osmotic depletion with special reference to the formation of compound words and their derivatives in English.

As we shall see, the depletion phenomenon under discussion here is effected in a variety of ways. It may be realized by clipping, blending or attrition, which we will take up in order in the main body of our discussion. Our treatment of clipping will include both back clipping and fore clipping.

In our discussion of blending, we will deal with not just the typical blending phenomenon per se but also cluster simplification that often figures in blending. We will then conclude with a look at feature inheritance, cliticization and other phenomena of descriptive and explanatory relevance that stem from the kinds of structural compression addressed in the main body of our discussion.

## 2. Clipping

Clipping is often employed in trimming a supralexicalexpression down to size so as to fit it into the cramped space of a word. As already pointed out, both back clipping and fore clipping operate in the downsizing process under discussion here. As we shall soon see, however, the line of demarcation between the two types of clipping is sometimes blurred and thus anything but clear-cut.

### 2.1. Back Clipping

Let us begin by considering the derivational process exemplified by such (derivational) paraphrase pairs as those cited below.

- (1) a. Harvard      (<=Harvard University)
- b. Oxford      (<=Oxford University)
- c. Cambridge (<=Cambridge University)

Being a bilexical form, the second member of each paraphrase pair here may often be felt to be on the obese side for the skimpy slot it gets assigned as a noun. As a result, this bilexical form often ends up getting downsized to a monolexical form through the clipping of its head noun, i.e. *University*.

Note that the constituent that gets deleted here is a head word shared by all the supralexicalexpressions in question and by numerous other such forms in English as well. Thus this deleted constituent is arguably non-contrastive to the point of being affixal and thus relatively light of semantic content, which apparently also conduces to its eventual deletion. Incidentally, something like this seems to be true of virtually every underlying head constituent that gets deleted in the kind of downsizing addressed throughout our discussion.

It goes without saying that similar clipping figures in the derivations illustrated by paraphrase pairs such as the following.

- (2) a. the Pacific (<=the Pacific Ocean)  
 b. the Caribbean (<=the Caribbean Sea)
- (3) a. the Himalayas (<=the Himalaya Mountains)  
 b. the Philippines (<=the Philippine Islands)  
 c. the Balkans (<=the Balkan States)  
 d. the Olympics (<=the Olympic Games)

Note that in each of the derivations in (3) above the plural suffix *-s* gets transported from the deleted underlying head noun to the modifier that precedes it. This transportation phenomenon, universally applicable to the kind of head deletion under discussion here, will be gone into at greater length when we deal with feature inheritance later on.

An essentially identical account apparently applies to derivations such as those illustrated by the following paraphrase pairs.

- (4) a. class actions (<=class-action (law)suits)  
 b. J-bars (<=J-bar lifts)

It is interesting that a similar process is apparently at work in the derivation of the non-count noun in first member of each paraphrase set below from its count counterpart in the second and third members of the same set.

- (5) a. chicken (<=chicken meat/flesh (<=meat/flesh from a chicken))  
 b. lamb (<=lamb meat/flesh (<=meat/flesh from a lamb))  
 c. whale (<=whale meat/flesh (<=meat/flesh from a whale))  
 d. tuna (<=tuna meat/flesh (<=flesh/flesh from a tuna))

An inherently count noun here, e.g. *chicken*, derives its non-count characteristic from the inherently non-count head noun that gets deleted, i.e. *meat/flesh*. Thus the non-countness of the deleted underlying head noun here is bequeathed to the modifier that precedes it. We will also be revisiting this kind of transportation phenomenon in greater detail later on in connection with our discussion of feature inheritance.

Parenthetically, the clipping involved in (5) above may be either back or

fore clipping depending on which of the parenthesized expressions is taken as the underlying form. There does not appear to be any compelling evidence for or against either perspective here so that one is just as viable as the other. Mostly for ease of exposition, however, we are taking the first of the two parenthesized supralelexical forms as the underlying form for each of the words in question with the result that back clipping is assumedly involved here.

An exactly identical process of derivation may be assumed for numerous other non-count usages of inherently count nouns such as those illustrated in the derivations below.

- (6) a. oak           (<=oak wood           (<=wood from an oak))  
       b. mahogany (<=mahogany wood (<=wood from a mahogany))
- (7) a. mink       (<=mink fur       (<=fur from a mink))  
       b. fox       (<=fox fur       (<=fur from a fox))
- (8) a. crocodile (<=crocodile skin (<=skin from a crocodile))  
       b. alligator (<=alligator skin (<=skin from an alligator))

It is interesting to note at this point that our discussion provides a principled explanation for the metamorphosis of *mad cow* from count to non-count status, as illustrated in the paraphrase set below.

- (9) mad cow (<=mad cow disease (<=disease affecting a mad cow /  
       disease causing a cow to be mad))

Needless to say, paraphrase pairs such as the following go through a similar derivational process.

- (10) a. Alzheimer's (<=Alzheimer's disease)  
       b. Parkinson's (<=Parkinson's disease)
- (11) a. St. Paul's (<=St. Paul's Cathedral)  
       b. St. Peter's (<=St. Peter's Basilica)
- (12) a. a butcher's (<=a butcher's shop)  
       b. a florist's (<=a florist's shop)

We can offer a similar, clipping-based account for the morphing of non-count nouns into count nouns, as can be seen from an examination of paraphrase sets such as the following.

- (13) a. golds (←gold medals (←medals made of gold))  
 b. silvers (←silver medals (←medals made of silver))  
 c. bronzes (←bronze medals (←medals made of bronze))

The inherently non-count *gold*, *silver* and *bronze* above derive their countness from the count noun *medal*, which is the underlying head noun that gets deleted.

It goes without saying that an exactly identical phenomenon manifests itself in derivations such as those illustrated by the following paraphrase sets.

- (14) a. leathers (←leather clothes (←clothes made of leather))  
 b. silks (←silk shirts (←clothes made of silk))  
 c. rubbers (←rubbers overshoes (←overshoes made of rubber))

Incidentally, our account here sheds serendipitous light on such grammatically baffling temporal expressions as the following.

- (15) a. *a* twelvemonth (←*a* twelve-month *period*)  
 b. *a* fortnight (←*a* fortnight *period* / a fourteen-night *period*)

We assume that the derivations here involve the back clipping of *period* from the parenthesized underlying forms. The singularity of *twelvemonth* and *fortnight* here, as indicated by the indefinite article *a*, is then arguably inherited from the singularity of the deleted underlying head noun *period*. Thus the absence of the plural suffix *-s* from *twelvemonth* and *fortnight* here, despite the plurality of *twelve-* and *fort-* (←*fourteen-*), can quite naturally be explained along basically the same lines as the absence of the plural suffix *-s* from, say, *three-room* in *a three-room house*.

Back clipping of an essentially identical nature is apparently in evidence in the derivation of such de-adjectival nouns as those cited below.

- (16) a. valuables (←valuable *things*)  
 b. perishables (←perishable *things*)  
 c. (in)tangibles (←(in)tangible *things*)

- (17) a. *a* given           (<=*a* given *fact(or)/condition*)  
 b. *an* unknown       (<=*an* unknown *thing/person*)  
 c. *a* variable           (<=*a* variable *factor/quantity*)

It goes without saying that the number distinction and article predetermination displayed by the de-adjectival nouns here are inherited from the underlying head nouns that get deleted in the derivational process involved.

Back clipping also is arguably involved in the derivation of collective, de-adjectival nominals of the following form.

- (18) a. the rich   (<=*rich people*)  
 b. the poor   (<=*poor people*)

Note here that the human reference of either *the rich* or *the poor* arguably originates in *people*, the head noun that gets deleted from the underlying phrase. It may be in order here to make the point that a deleted underlying head noun always bequeaths its meaning to the form that remains after the deletion. The head noun *Mountain* deleted from *the Himalaya Mountains*, for example, leaves the semantic feature [+*mountain*] behind in *the Himalayas*.

Incidentally, we may just as well say that the derivation of either *the rich* or *the poor* here involves fore clipping rather than back clipping, for we can posit *people who are rich/poor* as an alternative underlying form. In the absence of any convincing evidence for either underlying-form candidate, however, our choice of back clipping over fore clipping here may admittedly be quite arbitrary.

Back clipping of just a part of the second constituent in a compound, rather than the whole thereof, is also observable in examples of relatively recent vintage, such as the following.

- (19) a. showbiz (<=*show business*)  
 b. webcam (<=*web camera*)

## 2.2. Fore Clipping

Fore clipping also figures quite prominently in the kind of structural condensation under discussion here, as can be seen from such paraphrase pairs as the following.

- (20) a. *the Hill* (<=*Capitol Hill*)  
 b. *the Street* (<=*Wall Street*)  
 c. *the Yard* (<=*Scotland Yard*)  
 d. *the Village* (<=*Greenwich Village*)

Parenthetically, the definite article *the* that figures in the derivations here is arguably a compensatory trace for the words deleted, i.e. *Capitol* in (20a), *Wall* in (20b), *Scotland* (20c) and *Greenwich* in (20d). Incidentally, similarly compensatory is the *the* that figures in *the rich* (<=*rich people*) and *the poor* (<=*poor people*) cited in (18) above. A detailed discussion of this kind of compensatory *the* is presented elsewhere (Park, 1992, 2001, 2002).

Needless to say, a basically identical account applies to such examples as the following, except that *the* here is not compensatory but underlyingly real.

- (21) a. *the Times* (<=*the New York Times*)  
 b. *the Lakers* (<=*the Los Angeles Lakers*)

Similarly motivated fore clipping arguably can often be observed in the transformation of proper nouns into common nouns, as can be seen from the examples below.

- (22) a. *two Picassos* (<=*two pictures by Picasso*)  
 b. *a Rembrandt* (<=*a picture by Rembrandt*)
- (23) a. *a Shakespeare* (<=*a playwright like Shakespeare*)  
 b. *Einsteins* (<=*scientists like Einstein*)

The assumption that fore clipping, rather than back clipping, figures here is admittedly rather arbitrary. For *two Picassos*, for one, may be thought of as originating in *two Picasso pictures*, in which case the derivation involves back clipping, not fore clipping. Once again, it is apparently open to debate which of the two alternatives we should adopt for our analysis here. Incidentally, we will have more to say about proper-to-common noun transformation when we deal with feature inheritance toward the end of our discussion.

Fore clipping also may apparently play a pivotal role in the derivation

of a denominal verb, as can be seen from the following examples.

- (24) a. *to single* (<=*to hit a single*)  
 b. *to double* (<=*to hit a double*)  
 c. *to triple* (<=*to hit a triple*)  
 d. *to homer* (<=*to hit a homer*)

Note here that the verb, *to single*, for one, owes its verbal characteristics (such as its inflectional range) to the deleted underlying verb *hit*. More examples of conversion like those cited here will be considered later on in our discussion of feature inheritance.

Sometimes part of a word, rather than the whole thereof, is what gets deleted in the kind of fore clipping in question here, as can be seen from the following paraphrase pairs.

- (25) a. *indoor* (<=*within (the) door*)  
 b. *drawing room* (<=*withdrawing room*)

From the history of English, we know that the the first member of each pair here originates in the second member of the same pair with the prefixal *with-* getting suppressed in the derivation of the former from the latter. Incidentally, it may very well be that *stood* (as in *Our friendship has stood the test of time*) originates in *withstood* (as in *Our friendship has withstood the test of time*).

Among similar examples of more recent vintage are the following.

- (26) a. 'shroomer (<=*mushroomer*)  
 b. blog/'blog (<=*weblog*)  
 c. Net/'Net (<=*Internet*)

We may note at this point that the derivation of a bahuvrihi compound may involve fore clipping, as can be seen from paraphrase pairs such as the following.

- (27) a. *a butterfingers* (<=*a person with butterfingers*)  
 b. *a lazybones* (<=*a person with lazy bones*)  
 c. *a loudmouth* (<=*a person with a loud mouth*)  
 d. *a redcap* (<=*a person wearing a red cap*)

We might just as well posit back clipping here if we assume that *redcap*, for one, derives from *red-capped person*. Parenthetically, the semantic feature [+human] associated with each of the bahuvrihi compounds here is due to the deleted underlying head noun *person*, which is an example of feature inheritance to be discussed in detail later on.

### 3. Blending

Blending also is arguably designed to help squeeze a supralexical underlying form into the inherently restricted space of a word. We will begin our discussion here by dealing with typical examples of blending such as *motel* (<=*motor* + *hotel*). We will then turn our attention to cluster simplification, which often occurs in the blending process. A particularly interesting type of cluster simplification is word-cluster simplification, which is effected by the deletion of the medial word from a three-word cluster, as in the derivation of *newsboy* from *newspaper boy*.

#### 3.1. Typical Blending

In most instances, blending typically involves syncopation that clips the middle portion of a supralexical form when it morphs into a word. Here the clipped portion is normally part of either the back of the first constituent or the front of the second constituent or both. The constituent that gets partially clipped in the blending process here tends to get (orthographically) cliticized and thus attached to the other constituent(s) with which it is in construction.

Let us begin our discussion here with a look at syncopation as it relates to the formation of the following blends.

- (28) a. smog            (<=*smoke* + *fog*)  
       b. telecast       (<=*television* + *broadcast*)  
       c. simulcast    (<=*simultaneous* + *broadcast*)  
       d. brunch        (<=*breakfast* + *lunch*)  
       e. bionics        (<=*biology* + *electronics*)  
       f. bit            (<=*binary* + *digit*)  
       g. magalog      (<=*magazine* + *catalog*)

In each of the derivations here, syncope clips the back of the first underlying constituent word and the front of the second underlying constituent word.

Syncope can clip just the front of the second underlying constituent word, as shown in each of the derivations below.

- (29) a. breathalizer (<=breath *analyzer*)  
 b. fanzine (<=fan *magazine*)  
 c. webzine (<=web *magazine*)  
 d. travelogue (<=travel *monologue*)

Syncope can also affect just the back of the first underlying constituent word, as shown in each of the derivations below.

- (30) a. paratrooper (<=parachute trooper)  
 b. heliport (<=helicopter port)  
 c. fortnight (<=fourteen night)  
 d. prep school (<=preparatory school)

It is interesting that the normally diphthongal nucleus of *-night* in (30c) may sometimes be abraded to the monophthongal /i/, which may count as an instance of attrition that we will be dealing with later on.

The kind of depletion involved in (30) above also may figure in the formation of certain prefixes. As a case in point, we may cite the recent emergence of the prefix *e-* illustrated in the paraphrase pairs below.

- (31) a. e-mail (<=E-mail (<=*electronic* mail))  
 b. e-commerce (<=*electronic* commerce)  
 c. e-money (<=*electronic* money)  
 d. e-book (<=*electronic* book)

The adjective *electronic* gets whittled down to the prefix *e-* here so as to arguably help the resulting form better accommodate itself into its cramped space as a lexical item. Note in this connection that *electronic* is common to all the supralexic forms cited above so that it is non-contrastive to the point of being affixal and thus light of semantic content. This also appears to play a role in the evolutionary reduction of *electronic* to a mere prefix *e-* here. Along the lines of Park (1992, 2003), we may

argue that the status of *electronic* as a premodifier also conduces to its “demotion” to a suffix. We will return to this suffix later on in connection with our discussion of capitalization and boundary reduction.

Similarly rationalizable is the formation of military-related prefixes illustrated below.

- (32) a. M-1 (<=*Military (Equipment)* 1)  
 b. M-16 (<=*Military (Equipment)* 16)

- (33) a. B-17 (<=*Bomber* 17)  
 b. B-29 (<=*Bomber* 29)

- (34) a. F-14 (<=*Fighter* 14)  
 b. F-105 (<=*Fighter* 105)

Additional examples such as the following may be cited in support of our contention here.

- (35) a. The USS Arizona (<=*The United States Ship* Arizona)  
 b. The USS Saratoga (<=*The United States Ship* Saratoga)

- (36) a. HMS Warrior (<=*Her Majesty's Ship* Warrior)  
 b. HMS Winston Churchill (<=*Her Majesty's Ship* Winston Churchill)

Should *The USS* and *HMS* be assigned prefixal status, which is suggested here, so may such abbreviated titles as *Mr.*, *Mrs.*, *Dr.* and *Mt.* Note parenthetically that the blending that figures in (35) and (36) is multiple blending, comprising two or more drastically clipped elements of the sort typically involved in acronymization, which we will be considering a few paragraphs later.

Our discussion here apparently casts serendipitous light on the derivational link between the (quasi-)prefix *al-* and the intensifier *all*, which is its underlying source, as can be seen from derivations such as those illustrated below.

- (37) a. *alone* (<=*all one*)  
 b. *altogether* (<=*all together*)  
 c. *almighty* (<=*all mighty*)  
 d. *also* (<=*all so*)  
 e. *already* (<=*all ready*)  
 f. *almost* (<=*all most*)

We can say that the intensifier *all* has been snipped down to the prefix *a-* so as to render the resulting word osmotically fit, so to speak.

Let us now turn our attention to examples of blending involving double back clipping, as illustrated by the derivations below.

- (38) a. *sitcom* (<=*situation comedy*)  
 b. *telex* (<=*teleprinter exchange*)  
 c. *pro-am* (<=*professional-amateur*)  
 d. *interpol* (<=*International (Criminal) Police (Organization)*)  
 e. *ob-gyn* (<=*obstetrics and gynecology*)

Blending may often feature two or more maximally back-clipped constituents, as in the derivation of acronyms from their supralexic sources. Let's take the following acronymic derivations, for example.

- (39) a. *MC* (<=*master of ceremonies*)  
 b. *KO* (<=*knock out*)  
 c. *CPA* (<=*certified public accountant*)  
 d. *PDA* (<=*personal data assistant*)  
 e. *SIDS* (<=*sudden infant death syndrome*)  
 f. *BYOB* (<=*bring your own bottle*)

Needless to say, the examples of multiple clipping in (38) and (39) are arguably all osmotically motivated. Parenthetically, it is argued elsewhere (Park, 1993, 2002) that the capitalization of the initials that figure in acronyms, as in the examples here, is a hierarchical compensation for the drastic linear surgery involved in acronymization.

### 3.2. Cluster Simplification

Syncopation that figures in blending is often realized by declustering,

which typically takes the form of the medial member of a three-member cluster getting squeezed out in its entirety, under pressure from the two other members of the same cluster. For convenience of exposition, declustering here may be thought of as comprising three different types, i.e. (1) consonant-cluster simplification, (2) morpheme-cluster simplification, and (3) word-cluster simplification, which we will consider in order.

### 3.2.1. Consonant-Cluster Simplification

Let us begin by considering the following examples.

- (40) a. *Christmas* (<=Christ + -mas)  
 b. *chestnut* (<=chest + nut)  
 c. *mustn't* (<=must + -n't)

The consonant /t/, the medial member of the consonant cluster in each example here, gets erased under pressure from the consonant /s/ that precedes and the nasal consonant that follows.

Incidentally, the disappearance of the etymological /t/ in *Essex* (<=*East Saxon*) and *Wessex* (<=*West Saxon*) also is due apparently to the kind of consonant-cluster simplification under discussion here. It is interesting that *crackpot* originated in *cracked pot*, which clearly indicates that consonant-cluster simplification played a role in its formation. It is just as interesting that *hard-shell clam* is generally favored over *hard-shelled clam* probably because consonant-cluster simplification kicks in here. Park (1984) makes a fairly wide-ranging inquiry into consonant cluster simplification in English.

We can rationalize along similar lines the deletion of the etymological /d/ in words such as the following.

- (41) a. *handsome*  
 b. *Newfoundlander*

The deletion of the “th” sound in the following words affords us an additional example of the kind of consonant-cluster simplification under discussion here.

- (42) a. *nor'wester* (<=*northwester*)  
 b. *sou'wester* (<=*southwester*)

Incidentally, /r/ and /w/ are functionally consonantal so that they exert pressure on the intervening “*th*” sound to get deleted. Parenthetically, it may be on the analogy of the “*th*”-deletion here that we get *nor'easter* from *northeaster*, in which “*th*” is followed by a vowel, rather than a consonant.

Our discussion here seems to throw light on the disappearance of the *going*-final /g/ and the *used*-final /t/ in (43a) and (43b), respectively.

- (43) a. I'm *gonna* visit New York.  
 (<=I'm *going to* [gowingtʊ] visit New York.)  
 b. I *used to* [yustu] work a lot harder.  
 (<=I *used to* [yusttu]work a lot harder.)

We are assuming here that the underlying consonant cluster in question is something like /ngt/ in (43a) and /stt/ in (43b), both of which are historically real.

Similarly motivated declustering manifests itself on the orthographic plane as well, as is attested to by examples such as the following.

- (44) a. *Firland* (<=Finn + land)  
 b. *Lapland* (<=Lapp + land)  
 c. *Milton* (<=Mill + -ton)  
 d. *pastime* (<=pass + time)  
 e. *welcome* (<=well + come)

The underlying three-consonant cluster in question in each derivation above arguably is phonetically real, in which case the cluster simplification here is not just orthographic but also phonological. Something like this point is made in Park (2002).

Similar orthographic declustering is observable in the optional deletion of the parenthesized tokens of *l* in the following examples.

- (45) a. *wil(l)ful* (<=will + full)  
 b. *skil(l)ful* (<=skill + full)

Incidentally, the reduction of *full* to *-ful* here will be discussed later on in connection with the attrition phenomenon.

### 3.2.2. Morpheme-Cluster Simplification

Supposing that the first member of each paraphrase pair below derives from the second member of the same pair, we arguably have a case of morpheme-cluster simplification here.

- (46) a. He talks *as though* he is somebody.  
 b. \*He takes *as although* he is somebody.
- (47) a. He's happy *even though* he's penniless.  
 b. \*He's happy *even although* he's penniless.

As in consonant-cluster simplification, the medial member of the three-member cluster in either case here, i.e. *al-*, gets squeezed out under pressure from its two neighbours, i.e. *as/even* in front and *though* in back. Needless to say, this squeezing out is arguably motivated to help ease the supralexical subordinate conjunction in question here into its rather space-strapped slot as a lexical item.

The deletion of the morpheme *a-* from *around* in the examples below is arguably amenable to an essentially identical interpretation.

- (48) a. all-round ( $\leftarrow$ =all-around)  
 b. all-rounder ( $\leftarrow$ =?all-arounder)

Note that the deletion of the morpheme *a-* here is optional in (48a) and obligatory in (48b). Along the lines of Park (2003), we may argue that *all-around-* in *?all-arounder* premodifies the (agent) "nominal" *-er* so that the morpheme *a-* in *?all-arounder* is under added pressure to get deleted so as to help lighten up the compound premodifier. Thus *a-*, the medial member of the morpheme cluster here, is arguably under greater pressure to get forced out in (48b) than in (48a), which is evidently why its deletion is obligatory in (48b) and only optional in (48a).

### 3.2.3. Word-Cluster Simplification

Word-cluster simplification appears to be fairly common in compound-word formation in English, as is attested to by examples such as the following.

- (49) a. wastebasket (<=wastepaper basket)  
 b. newsboy (<=newspaper boy)  
 c. water dog (<=waterfowl dog)  
 d. twelfth tide (<=twelfth *night/day* tide)  
 e. twelfth-cake (<=twelfth *night* cake)

As in the two other types of declustering already examined, the medial member of the cluster in question in each derivation above ends up getting deleted so as for the resulting word to arguably better fit into its space-strapped slot as a lexical item.

We may observe in this connection that word-cluster simplification frequently involves deletion of references to people, as can be seen from an examination of examples such as the following.

- (50) a. madhouse (<=mad *people's* house (<=house for mad *people*))  
 b. poorhouse (<=poor *people's* house (<=house for poor *people*))  
 c. sickroom (<=sick *person's* room (<=room for a sick *person*))  
 d. sickbed (<=sick *person's* bed (<=bed for a sick *person*))  
 e. sick leave (<=sick *person's* leave (<=leave for a sick *person*))  
 f. disabled rights/access (<=disabled *people's* rights/access (<=rights/access of disabled *people*))  
 g. mental/psychiatric hospital (<=mental/psychiatric *patients'* hospital (<=hospital for mental/psychiatric *patients*))

Just as frequently deleted in word-cluster simplification are non-human references, as can be seen from examples such as the following.

- (51) a. New Church  
 (<=New *Jerusalem* Church (<=Church of the New *Jerusalem*))  
 b. foreign minister  
 (<=foreign *affairs* minister (<=minister of foreign *affairs*))  
 c. majority leader  
 (<=majority *party* leader (<=leader of the majority *party*))  
 d. hard/soft clam  
 (<=hard/soft-*shell* clam (<=clam with a hard/soft *shell*))  
 e. rape oil (<=rapeseed oil (<=oil from rapeseed))  
 f. fly rod (<=fly-*fishing* rod (<=rod for fly *fishing*))

Additional examples of a similar nature such as the following may be cited almost *ad nauseum*.

- (52) a. white sale (<=white-goods sale (<=sale of white goods))  
 b. dry measure  
     (<=dry-commodities measure (<=measure for dry commodities))  
 c. white book (<=white-bound book (<=book with white covers))  
 d. blue book (<=blue-covered book (<=book with blue covers))  
 e. green card (<=green-colored card (<=card in green color))  
 f. greenhouse (<=green-plant(s) house (<=house for green plants))  
 g. criminal court (<=criminal-case(s) court (<=court for criminal cases))

It may at this point be noted, as a matter of general principle, that what gets deleted in word-cluster simplification is semantically real in that its meaning figures in one way or another in the semantic interpretation of the compound word in question. We will have occasion to revisit this matter later on as we deal with feature inheritance.

The following example of word-cluster simplification is rather unique in that what gets deleted is not just the medial member of the underlying word cluster but also the back parts of the first and last members of the same cluster.

- (53) biopic (<=biographical motion picture)

Were the underlying form here posited as just *biographic picture*, then the derivation of *biopic* would belong in 2.1., i.e. in a discussion of typical blending.

Parenthetically, word-cluster simplification apparently affords us a principled account for the baffling absence of verbs of locomotion from expressions such as the following.

- (54) a. Murder *will out*. (=Murder *will come out*.)  
 b. Let me *down*. (=Let me *get down*.)  
 c. Help me *up*. (=Help me *get up*.)

Positing an underlying verb of locomotion for each of the expressions cited here is arguably justifiable on two counts at least. That is, it would be instrumental in assigning an intuitively correct semantic interpretation

to the expression in question and in simultaneously resolving the anomaly of a predicate lacking a bona fide main verb. The underlying main verb (of locomotion) here such as *come* or *get* may apparently get squeezed out under pressure from the preceding auxiliary (such as *will*, *let* or *help*) and the following adverbial particle (such as *out*, *down* or *up*).

Our perspective here seems to shed interesting light on the deletion, especially in American English, of the preposition *of* in derivations such as those cited below.

- (55) a. Look *out the* window. (=<Look *out of the* window.)  
 b. She put the cat *out the* door. (=<She put the cat *out of the* door.)  
 c. Don't throw butts *out the* car window.  
 (=Don't butts *out of the* car window.)  
 d. Beer was flying *out the* door. (=<Beer was flying *out of the* door.)  
 e. People poured *out the* stadium. (=<People poured *out of the* stadium.)  
 f. A guy called *out the* window. (=<A guy called *out of the* window.)

We are making the admittedly controversial assumption that *out the* here as well as *will out*, *let down* and *help up* in (54) are compound words of sorts.

Incidentally, we may be justified in invoking word-cluster simplification in explaining the inadmissibility of *have got to* in the sense of “*must*” after auxiliary verbs such as *will* and *may*, as opposed to the admissibility of its allo-form *have to* in the same environment. The auxiliary verbs here may help compress the periphrastic equivalent of *must* in question here into a far more close-knit cluster of words than it would otherwise be the case. Should this be the case, then *got* in *have got to* here may arguably be under especially heavy pressure to drop out after an auxiliary verb.

The kind of word-cluster simplification under discussion here is arguably involved in the derivation of such place names as those cited below.

- (56) a. Stratford-(up)on-Avon (=<Stratford-(up)on-*the-River*-Avon)  
 b. Newcastle-upon-Tyne (=<Newcastle-upon-*the-River*-Tyne)  
 c. Annandale-on-Hudson (=<Annandale-on-*the-River*-Hudson)

In connection with (56) above, it may be observed that what gets deleted is not only the content word *River* but also the function word *the*

preceding it. In fact, the definite article and other function words such as prepositions may often get deleted as part of word-cluster simplification. Let us take derivations such as the following, for example.

- (57) a. midweek ( $\leftarrow$ (*the*) *middle of the week*)  
 b. midnight ( $\leftarrow$ (*the*) *middle of the night*)  
 c. midstream ( $\leftarrow$ (*the*) *middle of the stream*)  
 d. midair ( $\leftarrow$ (*the*) *middle of the air*)

Note that the cluster simplification here involves the deletion of the preposition *of* and the definite article *the* as well as *-dle*, which is a portion of the content word *middle*.

We may argue that the preposition *in* also often gets deleted in word-cluster simplification, as can be seen from examples such as the following.

- (58) a. Korean-made cars ( $\leftarrow$ (cars made *in* Korea)  
 b. an American educated scientist ( $\leftarrow$ (scientist educated *in* America)  
 c. a Newsweek article ( $\leftarrow$ (an article *in* Newsweek)

We also may quite plausibly argue that the word-formation process involved in the derivation of words like the following goes through a word-cluster simplification that features the deletion of the generic *the*.

- (59) a. uphill ( $\leftarrow$ (up *the* hill)  
 b. downstream ( $\leftarrow$ (down *the* stream)  
 c. underground ( $\leftarrow$ (under *the* ground)  
 d. overhead ( $\leftarrow$ (over *the* head)  
 g. onsite ( $\leftarrow$ (on *the* site)  
 f. offshore ( $\leftarrow$ (off *the* shore)  
 g. inhouse ( $\leftarrow$ (in *the* house)

Along the lines of Park (2003), we may argue that the deletion of *the* here may have further been facilitated by the need for the resulting words in question to shed weight because of their frequent use as premodifiers.

It is interesting that the assumed historical derivation cited below also exemplifies the kind of word cluster simplification under discussion here.

- (60) because ( $\leftarrow$ (by cause ( $\leftarrow$ (by *the* cause)))

The assumed three-word cluster *by the cause* apparently got simplified to *by cause* (and eventually to *because*) by losing the medial word *the*, probably well before Middle English times.

Note incidentally that word-cluster simplification of a similar sort apparently played a role in the historical derivation of the first member of each paraphrase pair below from the second member of the same pair.

- (61) a. at least (<=*at the* least)  
 b. at best (<=*at the* best)  
 c. at last (<=*at the* last)  
 d. at table (<=*at the* table)

Incidentally, it seems that the idiomatic expressions cited here regularly sported *the* until well into early Modern English times.

The disappearance of *the* in the derivation of words such as the following may be accounted for along essentially identical lines.

- (62) a. aboard (<=*on board* (<=*on the board* (of)))  
 b. ashore (<=*on shore* (<=*on the shore*))

Incidentally, the deletion of /n/ from the underlying preposition *on* here is similar in motivation to the deletion of /n/ from the indefinite article *an* in the formation of its allomorph *a*. This phenomenon is treated in some detail elsewhere, e.g. in Park (1984).

Our discussion here may help illuminate the rather curious absence of the article, either definite or indefinite, from the noun constituents of idiomatic expressions such as those in italics below.

- (63) a. The woman *with child* was treated first.  
 b. She did it *on foot*.  
 c. *Take note* of the road conditions.  
 c. He dismissed my idea *out of hand*.  
 d. Elections will now *take place* on November 25.  
 e. The government was once more *out of step* with public opinion.  
 f. He knows hundreds of poems *by heart*.  
 g. Integrity is the word that *comes to mind* when talking of the man.  
 h. I just *played* it *by ear*.

An idiom being a kind of *compound word by definition*, we may be justified in invoking word-cluster simplification to explain the absence of *the* or *a(n)* from in front of the nominal constituent in each of the idioms here.

We may, with some justification, include the following examples in our discussion of word-cluster simplification.

- (64) a. B.A. (<=bachelor *of* arts)  
 b. MC (<=master *of* ceremonies)
- (65) a. ob-gyn (<=obstetrics *and* gynecology)  
 b. ENT (<=ear, nose *and* throat)

The deletion of *of* and *and* here, along with the drastic depletion of the constituents they link together, is arguably designed to help the supralexicial underlying forms here morph more comfortably into their respective surface forms as lexical items.

#### 4. Attrition

A compound word often gets worn down to size in other ways than by *clipping and blending*. Also apparently motivated to help fit the resulting word into its space-strapped slot as a lexical item, attrition often takes the form of an abraded vocalic nucleus in the semantically lesser of the two constituent words making up the supralexicial underlying form in question. It may also take the form of suffixization, as when the semantically lesser of the two constituent words in question gets reduced to *-er* or *-s*.

##### 4.1. Nucleic Attrition

Let us consider the following compound words with special reference to the pronunciation of *-mouth*.

- (66) a. *Dartmouth*  
 b. *Plymouth*  
 c. *Portsmouth*  
 d. *Monmouth*  
 e. *Tynemouth*

All the words here terminate with *-mouth* so that it is non-contrastive to the point of being suffixal and thus relatively light of semantic content. In conjunction with the need for each of the compound forms here to ease into its restricted space as a lexical item, this arguably helps reduce the originally diphthongal nucleus of *-mouth* to a weak monophthong, i.e. a mere schwa.

An identical perspective apparently sheds light on the attrition of the vocalic nucleus in the second constituent of each of the compound words cited below.

- (67) a. *England*  
 b. *Scotland*  
 c. *Ireland*  
 d. *Finland*  
 e. *Iceland*  
 f. *Poland*

- (68) a. *Englishman*  
 b. *Frenchman*  
 c. *Irishman*  
 d. *Welshman*  
 e. *Scotchman/Scotsman*

- (69) a. *Sunday*  
 b. *Monday*  
 c. *Tuesday*  
 d. *Wednesday*  
 e. *Thursday*  
 f. *Friday*  
 g. *Saturday*

Vocalic-nucleus attrition may well be at work, albeit to a lesser degree, in the derivation of each of the compounds below from its phrasal underlying form given in parentheses.

- (70) a. *greenbug* (<=green *bug*)  
 b. *blackbird* (<=black *bird*)  
 c. *bluefish* (<=blue *fish*)

In each pair here, the second constituent in italics gets considerably weaker stress and thus a phonetically weaker vocalic nucleus in the compound than in the phrasal combination.

Incidentally, assigning "suffixal" status to the second constituent of each and every compound cited so far in 3.1. seems to be fully justifiable, especially in light of the historical derivation of such suffixes as *-ly* and *-ful*, as illustrated below.

- (71) a. *manly* (<=man + *like*)  
 b. *careful* (<=care + *full*)

Both *-ly* and *-ful*, which originate in independent words, have become fairly productive suffixes in present-day English apparently through a process of gradual attrition. Note in this connection that we witness here instances of not just nucleic attrition but also consonantal attrition, phonetically, if not phonemically.

Similar in motivation may be the attrition of *and* to /n/ (or to a schwa followed by /n/) in derivations such as the following.

- (72) a. *cut-and-dried* (<=cut *and* dried)  
 b. *hit-and-run* (<=hit *and* run)  
 c. *bread-and-butter* (<=bread *and* butter)  
 d. *hide-and-see*k (<=hide *and* seek)  
 e. *hard-and-fast* (<=hard *and* fast)  
 f. *bed-and-breakfast* (<=bed *and* breakfast)

Incidentally, /n/ here may be regarded as something of an infix, probably on the order of the infix /n/ in words such as *messennger* (from *message*) and *passennger* (from *passage*).

A special case of attrition seems to be afforded by certain numerical expressions in ordinary prose, such as that which one may encounter in a journalistic piece. Let's take the following paraphrase pairs, for example.

- (73) a. *17* hours (<=*seventeen* hours)  
 b. *69* hours (<=*sixty-nine* hours)  
 c. *90* hours (<=*ninety* hours)

- (74) a. *seven* hours (<=7 hours)  
 b. *nine* hours (<=9 hours)  
 c. *zero* hours (<=0 hours)

It appears that of the two equivalent expressions in each paraphrase pair here, the first one tends to be preferred to the second one in day-to-day prose. Should this indeed be the case, we can say that Arabic numerals are apparently favored over their verbal equivalents for compound numbers and vice versa for simple numbers. Then numerals arguably being lighter versions of their verbal equivalents, the proclivity for Arabic numerals for compound numbers may very well be explained as resulting from the kind of attrition under discussion here.

The attrition phenomenon under discussion here also is apparently in evidence in the derivation of the following numeral compounds of recent vintage.

- (75) a. 9/11 (<=September 11)  
 b. 24/7 (<=24 hours a day, seven days a week)

### 3.2. Suffixization

The kind of attrition under discussion here seems to help illuminate the origins of the suffix *-er* in some of its manifestations, as in the derivations below.

- (76) a. *diner* (<= *dining car*)  
 b. *sleeper* (<= *sleeping car*)  
 c. *smoker* (<= *smoking car*)
- (77) a. *washer* (<= *washing machine*)  
 b. *dryer* (<= *drying machine*)

Being non-contrastive to the extent of being suffixal and thus semantically lightweight, *-ing car/machine* above arguably gets worn down to the suffix *-er*, thereby facilitating the resulting words' accommodation in their space-strapped slots as lexical items.

Needless to say, an exactly identical explanation apparently applies to the role of attrition in the formation of the suffix *-er* illustrated in the

examples below.

- (78) a. broiler (<= broiling *chicken*)  
 b. fryer (<= frying *chicken*)  
 c. roaster (<=roasting *chicken*)
- (79) a. cooker (<=cooking *appliance*)  
 b. blender (<=blending *appliance*)  
 c. toaster (<=toasting *appliance*)

The *-er* suffix in question here follows not just verbal bases but also inherently nominal bases, as can be seen from examples such as the following.

- (80) a. fiver (<=five-dollar bill / five-pound note)  
 b. tenner (<=ten-dollar bill / ten-pound note)  
 c. homer (<=home run)  
 d. topper (<=top hat / top coat)  
 e. grounder (<=ground ball)

The kind of attrition under discussion here may sometimes lead to the rise of other suffixes than *-er*, such as quasi-plural suffix *-s* in the following derivations.

- (81) a. the NCAAs (<=the NCAA *Tournament*)  
 b. Dow Jones industrials (<=Dow Jones industrial *average*)  
 c. undies (<=underwear)  
 d. tails (<=tailcoat)  
 e. oats (<=oatmeal)

The kind of quasi-plural suffix illustrated here is discussed in more detail in Park (1992, 2002).

## 5. Related Phenomena

Related to (end)osmotic depletion under discussion are a number of interesting phenomena of relevance to the structural description and

explanation of English. Among these are feature inheritance, cliticization, and capitalization and boundary reduction, which we will take up in order.

### 5.1. Feature Inheritance

We have already seen that a word whose derivation involves deletion of the head of a supralexicical underlying form inherits features from that deleted head. Let's take the following, for example.

- (82) a. turkey (<=turkey *meat*, i.e. *meat from a turkey*)  
 b. cod (<=cod *meat*, i.e. *meat from a cod*)  
 c. pine (<=pine *wood*, i.e. *wood from a pine*)

The inherently count noun *turkey*, for one, inherits the non-count characteristic from the deleted underlying head noun *meat*, which is inherently non-count, and thus gets reclassified as a non-count noun.

It is clear from (82) and other such examples that the deleted head of the supralexicical underlying form tends to leave a trace of one sort or another behind in the resulting word in question. It is interesting that the law of conservation, as discussed in Park (1992), is thus apparently alive and kicking in all of the examples cited here in support of our thesis about feature inheritance.

We have also seen that nouns like the following, which involve head deletion in their derivation, are subject to feature inheritance in the other direction, i.e. from count to non-count.

- (83) a. nickels (<=nickel *coins*, i.e. *coins made of nickel*)  
 b. coppers (<=copper *coins*, i.e. *coins made of gold*)  
 c. nylons (<=nylon *stockings*, i.e. *stockings made of nylon*)  
 d. patent leathers  
 (<=patent leather *shoes*, i.e. *shoes made of patent leather*)

In each derivation here, the resultant word, which is an inherently non-count noun, inherits the feature of countness from the deleted head of the supralexicical underlying form, which is an inherently count noun. Thus the inherently non-count noun *nickel*, for one, becomes a count noun here by inheriting the countness inherent in the deleted underlying

head noun *coin(s)*.

Similar transformation of nouns, i.e. from non-count to count status, is also exemplified by derivations such as the following.

- (84) a. two coffees            (<=two *kinds/portions of coffee*)  
       b. many kindnesses    (<=many *acts/kinds of kindness*)  
       c. *a* Tolstoy            (<=*a writer like Tolstoy*)  
       d. *another* Mozart    (<=*another composer like Mozart*)

Once again, the countness of the deleted underlying head noun is inherited in each derivation here by the resultant word, which is an inherently non-count noun.

Feature inheritance also sheds light on such apparent anomalies as are associated with following proper nouns.

- (85) a. *the* Philippines    (<=*the Philippine Islands*)  
       b. *the* Himalayas    (<=*the Himalaya Mountains*)  
       c. *the* Balkans        (<=*the Balkan States*)

*The Philippines* here, for one, is a proper noun and yet behaves as if it is a common noun in that it is in plural form and predetermined by the (definite) article. This is highly erratic in that a proper noun is normally in mutually exclusive distribution with number distinction and article predetermination, which are distinctive features of a common noun. We can resolve this apparent anomaly, however, by resorting to our theory of feature inheritance. We can simply say that both the plural suffix and the definite article in each proper noun cited above originate in the same features associated with the deleted head noun of its supralexicalexical underlying source.

A similar resolution is applicable to the apparent anomaly of the definite article *the* used with proper nouns, as in the following expressions.

- (86) a. *the* Mediterranean    (<=*the Mediterranean Sea*)  
       b. *the* Arctic            (<=*the Arctic Ocean*)  
       c. *the* Thames            (<=*the Thames River / the River Thames*)

The definite article *the* in front of each proper noun here originates in the definite article *the* associated with the deleted underlying head noun,

which is a common noun.

We can provide a similar account for the presence of the plural suffix *-s* in *Nikes* and *Reeboks* in the following derivations.

- (87) a. *Nikes* (<=*Nike shoes*)  
 b. *Reeboks* (<=*Reebok shoes*)

The plural suffix *-s* in question here is inherited from the underlying plural head noun *shoes*, which gets deleted in the derivational process involved.

Feature inheritance also is at work in conversion, as can be seen from derivations of de-adjectival nouns such as the following.

- (88) a. *untouchables* (<=*untouchable persons*)  
 b. *marrieds* (<=*married persons*)  
 c. *durables* (<=*durable goods*)
- (89) a. *an extraterrestrial* (<=*an extraterrestrial being*)  
 b. *a human* (<=*a human being*)  
 c. *a primitive* (<=*a primitive thing or person*)

Needless to say, the number distinction and article predetermination displayed by each of the de-adjectival nouns here are inherited from the deleted underlying head noun in question.

Feature inheritance is also observable in adjective-to-verb conversion, as can be seen from derivations of de-adjectival verbs such as the following.

- (90) a. *to quadruple* (<=*to make/become quadruple*)  
 b. *to calm* (<=*to make/become calm*)  
 c. *to dry* (<=*to make/become dry*)

Each de-adjectival verb here inherits its verbness, including its inflectional features, from the underlying (light) verb that gets deleted in the derivational process involved.

Needless to say, a denominal verb also inherits its verbness, including its inflectional behavior, from the underlying (light) verb that gets deleted, as can be seen from derivations such as the following.

- (91) a. *to bottle* (<=*to put into a bottle*)  
 b. *to corner* (<=*to put into a corner*)  
 c. *to pocket* (<=*to put into one's pocket*)
- (92) a. *to peel* (<=*to remove the peel from*)  
 b. *to skin* (<=*to remove the skin from*)  
 c. *to bone* (<=*to remove the bones from*)
- (93) a. *to oil* (<=*to supply with oil*)  
 c. *to water* (<=*to supply with water*)  
 c. *to house* (<=*to provide with a "house"*)

Feature inheritance seems to manifest itself somewhat erratically in the formation of acronyms, as can be seen from an examination of the following acronymic proper nouns.

- (94) a. NASA (<=*the National Aeronautics and Space Administration*)  
 b. NATO (<=*the North Atlantic Treaty Organization*)  
 c. OPEC (<=*the Organization of Petroleum Exporting Countries*)  
 d. UNESCO (<=*the United Nations Educational, Scientific and Cultural Organization*)  
 e. NAFTA (<=*the North American Free Trade Agreement*)

Unlike the underlying compound forms from which they are derived, the acronyms here appear to have become true proper nouns in that they do not normally allow predetermination by the definite article *the*. Thus in none of the derivations above is the definite article *the* that is associated with the (depleted) underlying head noun inherited by the acronym in question.

Note that acronymic proper nouns such as those cited in (94) above, which dispense with the underlying *the*, are normally read as words, not as sequences of letters. When acronymic proper nouns are read (not as words but) as sequences of letters, however, the underlying *the* may or may not be suppressed. It is anything but clear exactly what conditions the retention and suppression of this underlying *the* here.

An acronymic proper noun, pronounced as a sequence of letters, often dispenses with the underlying *the*, especially if it refers to a unique institution familiar to the interlocutors. Cases in point are afforded by

acronymic names of universities such as the following.

- (95) a. UCLA (<=*the* University of California at Los Angeles)  
 b. MIT (<=*the* Massachusetts Institute of Technology)  
 c. USC (<=*the* University of Southern California)  
 d. CUNY (<=*the* City University of New York)

An acronymic proper noun, even when read as a sequence of letters, often inherits and retains the underlying *the* in question here, however. Let's take the following, for example.

- (96) a. *the* FBI (<=*the* Federal Bureau of Investigation)  
 b. *the* FDA (<=*the* Food and Drug Administration)  
 c. *the* CIA (<=*the* Central Intelligence Agency)  
 d. *the* NBA (<=*the* National Basketball Association)  
 e. *the* WTO (<=*the* World Trade Organization)

To the extent that they inherit and retain the underlying *the*, the acronyms here may be kind of ambivalent between common and proper noun readings.

It is worth observing in this connection that *WHO* (short for *the World Health Organization*) may or may not inherit the underlying *the*, its inheritance tending to be less normal than its suppression. This may have to do with *WHO* being read more often as a word than as a sequence of letters.

Just as is the case with the relevant non-acronymic examples cited earlier in the current discussion, however, inflections apparently are consistently subject to inheritance in the derivation of acronymic verbs and nouns, as can be seen from the following examples.

- (97) a. p.o.'d (<=*pissed* off)  
 b. KO'd / KO'ing (<=*knocked* out / *knocking* out)  
 c. OD'd / OD'ing (<=*overdosed* / *overdosing*)
- (98) a. CEOs (<=*chief executive officers*)  
 b. the DJ's (<=*the disc jockey's*)  
 c. those MDs' (<=*those medical doctors'*)

So far our discussion has been confined to the inheritance of grammatical features. The deleted underlying head word bequeaths not just grammatical features but also semantic features to the resulting word. This inheritance of semantic features from the deleted underlying head word is in evidence in virtually all the examples cited in connection with our discussion of clipping and blending. Let's take the following clipping-based bahuvrihi compounds, for example.

- (99) a. a pot-belly ( $\leftarrow$ a *person with a pot belly*)  
 b. a birdbrain ( $\leftarrow$ a *person with a bird brain*)  
 c. a blockhead ( $\leftarrow$ a *person with a block head*)

The human reference of each of the bahuvrihi compounds cited here is arguably inherited from the head noun deleted from the underlying phrase. i.e. *person*. Needless to say, this kind of semantic inheritance from the deleted underlying head noun is crucial to the correct semantic interpretation of each of the expressions in question.

Semantic inheritance also may be illustrated with examples of word-cluster simplification such as the following.

- (100) a. sick bay ( $\leftarrow$ sick *people('s)* bay ( $\leftarrow$ bay *for sick people*)  
 b. poor farm ( $\leftarrow$ poor *people('s)* farm ( $\leftarrow$ farm *for poor people*)  
 c. insane asylum ( $\leftarrow$ insane *people('s)* asylum ( $\leftarrow$ asylum *for insane people*)

The human reference of each of the compound nouns here would not simply be retrievable were it not for the inheritance of the semantic feature [+human] from the deleted head noun *people('s)* in the first constituent of the supralexicial underlying form.

It seems that the source of semantic inheritance here does not have to be just the deleted underlying head word. Sometimes semantic features are apparently inherited from a deleted part of an underlying word, e.g. a prefix, as can be seen from the following examples.

- (101) a. drawing room ( $\leftarrow$ *withdrawing* room)  
 b. to stand (the test of time) ( $\leftarrow$ *to withstand* (the test of time))

Note here that *drawing room* and *to stand* apparently inherit their

respective semantic features of “*retreat/separation*” and “*resistance/ opposition*” from the deleted underlying prefix *with-*.

## 5.2. Cliticization

When a constituent of a word originating in a supralexic form gets depleted substantially enough, it apparently loses so much weight that it cannot stand by itself as an independent (orthographic) entity. As a result, this depleted constituent normally becomes a clitic and gets bound to the other word or words with which it is in immediate construction. This phenomenon is in evidence in derivations such as the following.

- (102) a. *H-bomb* (<=\**H* bomb (<=*hydrogen* bomb))  
 b. *A-bomb* (<=\**A* bomb (<=*atomic* bomb))  
 c. *N-bomb* (<=\**N* bomb (<=*neutron* bomb))

When all the constituents of a compound expression are radically depleted, as in acronyms, all the depleted constituents become clitics so that they end up clinging to each other for support, so to speak. Let us take the following acronyms, for example.

- (103) a. *MLB* (<=\**M L B* (<=*Major League Baseball*))  
 b. *CEO* (<=\**C E O* (<=*chief executive officer*))  
 c. *S.O.B.* (<=\**S. O. B.* (<=*son of a bitch*))

It appears that an essentially identical account is applicable to the cliticization illustrated by the examples below, where the sign & is a drastically depleted version of the coordinate conjunction *and*.

- (104) a. *P&G* (<=\**P & G* (<=*Proctor and Gamble / Proctor & Gamble*))  
 b. *S&L* (<=\**S & L* (<=*savings and loan*))

Note in this connection that not just & but also other signs/symbols such as \$ and % are all drastic abbreviations and thus apparently tend to be treated as (affixal) clitics in written English. As a result, they normally get attached to the numerals with which they are in construction, as in *\$100* and *100%*.

Incidentally, numerals may be quasi-clitics themselves arguably because

they are drastically abraded versions of their verbal equivalents. Should this really be the case, then the numeral *100* and the sign *\$/%* are bi-directionally bound to each other in *\$100* and *100%* here, instead of the latter being uni-directionally bound to the former. Recall our earlier comment on the difference in weight/mass between numerals and their verbal counterparts in connection with our discussion of the attrition phenomenon.

Incidentally, conjoined initials such as those cited in (104) above do not seem to behave like clitics when they are conjoined by the fully spelled-out version of *&*, i.e. *and*. Let us consider examples such as the following.

- (105) a. R&D / R and D (=<research and development)  
 b. R&R / R and R (=<rest and recreation)  
 c. B&B / B and B (=<bed and breakfast)  
 d. B&W / B and W (=<black and white)  
 e. A&M / A and M (=<Agricultural and Mechanical)

It is by no means clear why the two versions of the conjunction should behave so differently with respect to the cliticization of the initials they conjoin. Esthetics may possibly be a consideration here, as *RandD* would probably look more weird than *R and D*. Admittedly, however, this is a lame explanation at best.

It may be observed at this point that *&* does not seem to behave like a clitic by itself when it conjoins unabbreviated words. Thus *Proctor & Gamble* does not normally get written as *Proctor&Gamble*. It seems that abbreviated/initialized conjoins do not normally get bound to the unabbreviated *and* and vice versa. Along similar lines, we should be able to explain why *100%* and *one hundred percent* are okay, as opposed to the definite oddness of *100percent* and *one hundred%*.

As noted earlier, drastic depletion of other kinds such as that involved in blends of the more usual type also results in clitics, as in words such as the following.

- (106) a. updo (=<\*up do (=<upswept + hairdo))  
 b. mimmem (=<\*mim mem (=<mimicry + memorization))  
 c. Tex-Mex (=<\*Tex Mex (=<Texas + Mexico))

Note that every blend listed above comprises two constituents which are both drastically depleted versions of their respective underlying sources

and have thus both become clitics. Incidentally, the derivation of *mimmem* and *Tex-Mex* may arguably involve word-cluster simplification in that they may, quite plausibly, be thought of as deriving from *mimicry and memorization* and *Texas and Mexico*, respectively.

As we have shown in connection with examples like *H-bomb*, cliticization is often observed even when only one of the (two) constituents in a blended word is a drastically depleted version of its underlying source. We may cite the following additional examples, in each of which only the first underlying constituent gets drastically depleted and thus cliticized.

- (107) a. *Mideast* (<=\**Mid East* (<=*Middle East*))  
 b. *midfield* (<=\**mid field* (<=*middle field*))  
 c. *fortnight* (<=\**fort + night* (<=*fourteen night*))

The surname *Sinclair*, historically derived as shown below, is yet another example of a blend whose first constituent is a drastically depleted version of its underlying form and has thus become a clitic.

- (108) *Sinclair* (<=\**Sin Clair* (<=*Saint Clair*))

What motivates the depletion of *Saint* to *Sin-* in the above derivation is discussed in some detail elsewhere, e.g. in Park (1992, 2003).

Note in this connection that an abbreviated title (such as *Mrs.*) is a clitic that blends with what follows it to form a compound nominal, as in *Mrs. Johnson*. Although *Mrs.* is not orthographically bound to *Johnson* here, it is definitely a clitic in that it cannot occur independently of a name such as *Johnson*.

The following word affords us still another example of the first underlying constituent of a blend getting whittled down to a clitic.

- (109) *fulfill* (<=\**ful fill* (<=*full + fill*))

The second constituent here also may sometimes get worn down to a clitic with the result that both constituents get cliticized, giving rise to *fulfil*.

Sometimes blends may have only their second underlying constituents chipped down to clitics, e.g. the suffixes *-ful*, *-ly* and *-aholic*, as in the examples below.

- (110) a. *careful* (<=\*care *ful* (<=care + *full*))  
 b. *godly* (<=\*god *ly* (<=god + *like*))  
 c. *workaholic* (<=\*work *aholic* (<=work + (*alc*)*oholic*))

Parenthetically, the initial *a* in *-aholic* here replaces the initial *o* in *-oholic* so as to more faithfully reflect the weak vowel that it represents.

In each of the following examples of more recent origin also, (only) the second constituent of a supralexicical underlying form gets depleted to a clitic.

- (111) a. *showbiz* (<=show *business*)  
 b. *webcam* (<=web *camera*)
- (112) a. *webzine* (<=web *magazine*)  
 b. *fanzine* (<=fan *magazine*)

Sometimes deletion of *the* from a (nominal) second constituent in a supralexicical underlying form turns that constituent into something of a clitic that gets attached to the preposition, of which it is the complement. We may cite here compound words such as the following, in the derivation of each of which the underlying complement loses *the* and thus becomes a clitic that gets bound to the preposition in front of it.

- (113) a. *on-site* (<=\*on *site* (<=on + *the site*))  
 b. *offshore* (<=\*off *shore* (<=off + *the shore*))  
 c. *in-country* (<=\*in *country* (<=in + *the country*))

Sometimes a clitic results from the deletion of a head noun in the first constituent of a supralexicical underlying form, as in the examples below.

- (114) a. *wastebasket* (<=?waste *basket* (<=wastepaper + *basket*))  
 b. *sickbed* (<=?sick *bed* (<=sick *person's* + *bed*))  
 c. *poorhouse* (<=?poor *house* (<=poor *people's* + *house*))  
 d. *greenhouse* (<=?green *house* (<=green *plant(s)* + *house*))

In the derivation of *wastebasket* from *wastepaper basket*, for example, *waste* becomes a clitic when the head noun *paper* is deleted from the first underlying constituent *wastepaper*. Thus this cliticized *waste* winds

up getting bound to the second underlying constituent *basket*, giving rise to the ultimate surface form *wastebasket*.

Our argument here apparently breaks down in light of examples such as the following.

- (115) a. ?sickleave      (<=sick leave (<=sick *person's* + leave))  
       b. ?poorlaw        (<=poor law (<=poor *people's* + law))  
       c. ?greenthumb (<=green thumb (<=green *plant(s)* + thumb))

It may be speculated that *sick*, *poor* and *green* have had more time and/or more usage to become (orthographic) clitics in *sickbed*, *poorhouse* and *greenhouse* than in *sick leave*, *poor law* and *green thumb*, respectively. However, this does not seem to be all there is to it here. It may very well be that *sick*, *poor* and *green* in (115) above are clitics underlyingly although for one reason or another they are not yet treated as such orthographically.

### 5.3. Capitalization and Boundary Reduction

We have already noted that uppercase letters are employed to spell a typical acronym in its entirety, as in the following examples, arguably as an orthographic compensation for the portions radically deleted from its underlying form (Park, 1992, 2002).

- (116) a. OTC    (<=over the counter)  
       b. IPO    (<=initial public offering)  
       c. BTW   (<=by the way)

Acronyms sometimes employ lowercase letters either obligatorily or optionally, however, as can be seen from the following examples.

- (117) a. 5 kg  
       b. 60 rpm  
       c. 0.07 ppm
- (118) a. a.s.a.p. / A.S.A.P.  
       b. aka / AKA

The examples given in (117) may indicate that acronyms of a “suffixal”

nature tend to be written with lowercase letters. The examples given in (118) may indicate that acronyms of an adverbial nature sometimes allow their original uppercase forms to be lowercased. We may speculate that the lowercasing phenomenon here may have to do with the relatively light semantic content associated with the “suffixal” and adverbial acronyms in question. The obligatory lowercasing for the suffixal acronyms in (117), as opposed to the optional lowercasing for the adverbial acronyms in (118), may have to do with the fact that suffixes are generally less substantive than adverbs semantically. However, it is by no means clear exactly under what conditions the lowercasing is allowed or disallowed here.

It is interesting to note that the passage of time may sometimes have something to do with the decapitalization (or lowercasing) of an original uppercase initial, as is evidently the case with the example below.

(119) e-mail (<=E-mail (<=electronic mail))

When it first came into use around 1980, *electronic mail* was often shortened to *E-mail* with an uppercase initial in apparent compensation for the portion deleted from *electronic*, i.e. for *-lectronic*. Later on, the uppercase *E-* here seems to have gradually given way to its abraded version in the form of the lowercase *e-*, perhaps in recognition of its rapidly emerging status as a (semantically lightweight) prefix. Incidentally, the prefixal status of *e-* here must have been boosted by its widespread use in the 1980s and 1990s in not just *e-mail* but also a host of other newly minted *e-* words such as *e-commerce*, *e-text*, *e-money*, *e-book* and *e-government*.

Boundary reduction seems to be often resorted to in the formation of a compound or blended word to render the resulting word maximally fit as a lexical item. The following historical derivations afford us a case in point.

(120) a. maybe (<=may be)  
 b. because (<=by cause)

We know from the history of English that either word here used to be written in Middle English times as two orthographic words, i.e. as something like *may be* and *by cause*, respectively. The boundary between the two constituents in either word here has by now been so completely obliterated that today few of us realize that there was a time when it used to

orthographically separate the two constituent words in question.

The following example also may be cited from the recent history of the English language.

(121) email (<=e-mail (<= electronic mail))

Note here that *email* stems from *e-mail*, which in turn stems from *electronic mail*, in the course of which derivation the original orthographic boundary between the two constituents has gradually gotten reduced until there is no surface trace of it left in the latest form of the word, i.e. *email*.

The following derivation of relatively “recent” vintage affords us an additional example supportive of the point being made here.

(122) baseball (<=base-ball (<=base ball))

## References

- Jespersen, Otto. (1909-1949). *A Modern English Grammar on Historical Principles I-VII*. Copenhagen: Munksgaard.
- Nichols, W. R., Editorial Director. (2000). *Random House Webster's College Dictionary*. New York: Random House.
- Park, N-S. (1977). *Syntactic Variables in Readability with Special Reference to Korean EFL/ESL Learners*. Georgetown University Ph.D. Dissertation.
- Park, N-S. (1982). Constraint on proximate repetition with special reference to English. *Language Research* 18(1), 217-231.
- Park, N-S. (1984). The constraint on proximate repetition and phonological explanation in English. *Language Research* 20(4), 367-415.
- Park, N-S. (1992). The law of conservation and linguistic explanation. *Language Research* 28(2), 267-334.
- Park, N-S. (2002). Osmotic pressure and the structure of English. *Language Research* 38(2), 473-562.
- Park, N-S. (2003). Premodification and reduction in English. *Language Research* 39(4), 721-743.
- Quirk, R., S. Greenbaum, G. Leech, and J. Svartvik. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
- Summers, D., Director. (2000). *Longman Advanced American Dictionary*.

London: Longman.

Sinclair, J., Editor in Chief. (1995). *Collins Cobuild English Dictionary*.  
London: HarperCollins Publishers.

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. 1, 2004

Revised version received: May 15, 2004

Accepted: May 19, 2004