

## On Explication: Rationale, Method, Open Questions

**Rationale.** Explication is a type of analysis meant to unfold information in written words that is not just useful to describing and understanding written English, but also to teaching and learning it. Explication attempts to do so by identifying the basic units in written words – that is, the elements and particles.

**Elements and Particles.** Orthographic elements are the written counterpart of the morphemes of the spoken language. They are the smallest parts of written words that have the following two features:

- (i) In words in which they occur, they contribute the same or closely related semiotic, morphological, or syntactic content.
- (ii) They are spelled consistently from word to word – that is, spelled either the same or with variants that can be rationalized, as with the *+aesth* and *+esth* spellings in *aesthetics* and *esthetics*.

There are four kinds of elements: free bases like *plant* and *car*; bound bases like *fect* and *fer* as in *perfect* [per1+fect] and *defer* [de+fer1]; and prefixes and suffixes, like [re-] and -ed]1 in *repainted* [re+plant+ed]1.<sup>1</sup>

Elements often come in sets of two or more variants – for instance, *+arter*, *+arteri*, *+arterio* “windpipe, artery”, as in *artery*, *arterial*, *arteriosclerosis*. They also occur in what I call co-sets: pairs, or occasionally triplets, of elements that are related, but not necessarily equivalent, semiotically and that operate as a team – for instance, the co-set {*+ceed*, *+cess*} as in the verb *succeed* and its noun of result *success*.

Particles contribute no semiotic or syntactic content to a word, though they can have various functions. Particles enter in when elements are concatenated to form words. The most common particles are the second consonants that are inserted when a word like *run* takes a suffix that starts with a vowel: run1+n+ing]1 = *running*, the particle <*+n+*> serving the orthographic function of marking the preceding short vowel. Similar particles occur in the assimilation of prefix-final consonants – for instance, the <*+s+*> in *assimilation* [a/d+s+simil+at/e]1+ion]1, where the particle marks the easing of pronunciation. Other particles function as linkers, like the letter <*o*> common in technical words, especially from Greek: *ileostomy* ile1+o4+stom+y]3 and *ozonosphere* oz+on/e]3+o4+sphere.

**Arbitrariness and Motivation.** Explication speaks to Ferdinand de Saussure’s distinction between the arbitrariness and the relative motivation of the linguistic sign.

---

<sup>1</sup> In this article explications are underlined and are all from the Lexis database elsewhere on this site. Notice that although square brackets are commonly used to mark phonetic units, here they mark the beginning or end of written affixes. Letters and spellings are enclosed in pointed brackets.

Saussure divided the sign into the signifier (its expression) and the signified (its content). He makes his point unequivocally:

The bond between the signifier and the signified is arbitrary. Since I mean by the sign the whole that results from the associating of the signifier with the signified, I can simply say: *the linguistic sign is arbitrary*.<sup>2</sup> (His emphasis).

This much has been almost mantric in modern linguistics. However, Saussure went on to draw a distinction between this radical arbitrariness and a more orderly quality that he called *motivation*:

Some signs are absolutely arbitrary; in others we note, not its complete absence, but the presence of degrees of arbitrariness: *the sign may be relatively motivated*. (Saussure, 131, again his emphasis).

For example, a simplex word like, say, *six* is, in Saussure's view, absolutely arbitrary in its association of expression and content, as is evidenced by the fact that other languages have quite different expressions for conveying the content "six." However, a complex word like *sixteen* is not absolutely arbitrary and can be said to be relatively motivated because it can be analyzed into two components, *six* and *teen*, which he calls syntagms and I call elements. Each of these elements relates *sixteen* with several other words in the language: *Six* relates it paradigmatically to *sixty*, *sixth*, *twenty-six*, and so on. *Six* also relates it via scalar metonymy to *seven*, *seventy*, *seventh*, *twenty-seven* – and to all other ordinal and cardinal numbers in the number system. The element *teen* relates *sixteen* to such words as *thirteen*, *fourteen*, *teenage*, *fifteenth*, even *teener* and *teenybopper*. More remotely *teen* "10" relates *sixteen* to words like *thirty* and *fifty*, which contain the base *ty* "times ten". These paradigmatic relationships help provide the orderliness that Saussure calls relative motivation. Saussure says that "motivation varies, being always proportional to the ease of syntagmatic analysis and the obviousness of the meaning of the subunits present" (Saussure, 132). Explication is meant to increase "the ease of syntagmatic analysis" into elements and to heighten "the meaning [or I would say, the content] of the subunits."

Saussure also argues that

Everything that relates to language as a system must, I am convinced, be approached from this viewpoint, which has scarcely received the attention of linguists: the limiting of arbitrariness. . . . [T]he whole system of language is based on the irrational principle of the arbitrariness of the sign, which would lead to the worst sort of complication if applied without restriction. But the mind contrives to introduce a principle of order and regularity into certain parts of the mass of signs, and this is the role of relative motivation. (Saussure, 133)

---

<sup>2</sup>Ferdinand de Saussure, *Course in General Linguistics*, 1915. Repr. Charles Bally *et al*, eds. Wade Baskin trans. (NY: Philosophical Library, 1959) p. 67. His emphasis.

It is precisely these effects, “the limiting of arbitrariness” and the concomitant heightening of motivation, that explication works to increase in the written lexicon, as part of the search for increased order and regularity.

Explication attempts to analyze written words so as to highlight unifying strands of expression and content that link words into extended paradigms, both synchronic and diachronic. In this enterprise explication benefits from the fact that the English lexicon is “a small world”: Just as the social links are such that any two human beings on Earth are, we are told, separated by no more than six degrees of separation, so too the semiotic links of English are such that any two basic words in a lexicon of at least 30,000 types are separated by no more than three degrees of separation.<sup>3</sup> This tight linking and clustering is thought to reflect patterns of human cognition, and it leads to a mental lexicon organized so as to expedite the swift retrieval of words from memory. It appears that one of the functions (or at least one of the effects) of polysemy, other than simply reducing the number of word-types in the lexicon, is also to thicken that network of linkages. One of the goals of explication is to reveal relationships among elements so as to foreground these linkages and clusters.

## Desiderata:

**Simplicity.** Explication must heed Occam’s Razor – the proposition that in any analysis (i) entities should not be multiplied needlessly and (ii) the best analysis is the most economical and simple.

**Consistency.** One general corollary of Occam’s Razor is a principle of consistency. In logic and mathematics for an analysis to be consistent, it must contain no contradictions. For an explication to be consistent it should not contradict other explications of words of the same type in ways that can not be explained through more local rules and patterns. For instance, if you explicate, say, *communicative* as [com+munic+at/e]1+ive], it would be inconsistent and contradictory to explicate *noncommunicative* to [non+[com+mun/e+ic]1+at/e]1+ive] – for there are no known local rules or patterns to justify the use of munic in one but mun/e+ic]1 in the other. However it is not inconsistent to explicate *juror* to jur+or]2 but *conjuror* to [co/m+n+jur/e+or]2 since *conjuror* is derived from *conjure* with the silent final <e>, but there is no \*jure from which to derive *juror*.

**Ruliness.** By *ruliness* I mean the opposite of arbitrariness, marked with some degree of Saussure’s motivation. To the extent possible explication should be controlled by principles – or rules – that restrain arbitrary decisions. We need to learn enough about our orthographic system to have compelling reasons for decisions that so far I’ve made too often on the fly, too often quite arbitrarily. A total explication of the lexicon free of arbitrariness would be just that, total – though I suspect that is goal never to be realized. As Saussure said earlier, “Everything that relates to language as a system must . . . be approached from . . . : the limiting of arbitrariness.” Explication should add to that effort, at least so far as the system of the written lexicon is concerned .

---

<sup>3</sup> Adilson Motter, et al. “Topology of the Conceptual Network of Language,” *Physical Review E*, 65 065102@ (2002).

**Articulation.** Good articulation is having parts that are joined and fit together. Good articulation is central to the old Gestalt psychologists' notion of the good Gestalt. According to Kurt Koffka, "Roughly speaking, a minimum simplicity will be the simplicity of uniformity, a maximum simplicity that of perfect articulation".<sup>4</sup> Maximum simplicities produce Gestalten that are stable and useful; minimum simplicities do not. In explication a minimum simplicity is produced if any one of the four desiderata is followed too enthusiastically, underestimating or ignoring the importance of the other three. A maximum simplicity is produced if the four are all employed, are themselves articulated one with the other.

A reasonable analogy is a jigsaw puzzle in its various states: At the outset, with the pieces still in the box, the organization is utterly granular, with no articulation of any piece with any other. As the pieces are laid out on the table, right-side up, edge and corner pieces separated, other pieces grouped by color or partial image, there is some articulation as the relationship among pieces become more defined. As the puzzle is solved, piece by piece, things become increasingly articulated, until the completed puzzle represents perfect articulation, a fixed and stable and information-rich organization, free of arbitrariness. The process of explication should be much like that.

## Methods.

Some methodological corollaries of the desiderata above are (i) look first for the longest common string, (ii) examine definitions and etymologies, (iii) heed the principle of syllabicity, (iv) minimize complex procedures, (v) minimize vestiges, (vi) heed the principle of productivity, (vii) strive to avoid arbitrariness, and (viii) strive to present a plausible development of the word:

(i) Looking first for the longest common string is something like the search in arithmetic for the least common denominator. For instance, given a set of words to be explicated such as *collector*, *election*, *catalectic*, *intellectual*, *delectable*, *lecturer*, *lecterns*, *electuary*, *dialectal*, one's attention should be drawn first to the longest common letter string, <lect>. Then checking definitions and etymologies suggests a sorting into eight subgroups: words dealing (i) with gathering, (ii) with choosing, (iii) with leaving off, (iv) with perceiving, (v) with pleasing, (vi) with speaking, (vii) with reading, and – surprisingly – (viii) with licking:

gathering: *collector*  
choosing: *election*  
leaving off: *catalectic*  
perceiving: *intellectual*  
pleasing: *delectable*  
speaking: *dialectal*  
reading: *lecturer*, *lecterns*  
licking: *electuary*

---

<sup>4</sup> Kurt Koffka, *Principles of Gestalt Psychology*, NY: Harcourt, Brace, 1935, p. 171.

The eight subgroups vary so much in sense and usage that we are clearly dealing with some homographic elements and need to look more closely at earlier etymology, especially at any Proto-Indo-European (PIE) roots, which leads to these four groups<sup>5</sup>:

From PIE \*leg-1 “To collect”, later “to speak”: *collector, election, intellectual, dialectal, lecture, lectern*

From PIE \*slēg- “To be slack, languid”: *catalectic*,

From PIE \*leigh- “To lick”: *electuary*

Watkins gives no PIE source for *delectable*, but Pokorny, 673 proposes \*lēk2-, \*lək- “Branch, seize” with later senses of “snare, entrap.”<sup>6</sup>

This second grouping suggests that *catalectic, electuary, and delectable* contain at least three different homographic bases spelled <lect>. Whether the six words from PIE \*leg1- all contain the same base or more than one is, I think, a more subjective call, depending on whether one sees gathering, choosing, perceiving, speaking, and reading as close enough together in sense to all constitute one sense group – or two groups, or three, or four, or five, or six. Watkins maintains that the original PIE sense of \*leg1- was “to gather” with “to speak” as a derivative sense, suggesting that we are dealing here with two groups: (i) gather, choose, and (ii) speak, read – which is the explication proposed in the Lexis database:

lect1 “speak, read” (Total instances in Lexis: 46), as in:

acrolect	<u>acr1+o4+lect1</u>
dialect	<u>[dia+lect1</u>
idiolect	<u>idi+o4+lect1</u>
lectern	<u>lect1+ern]1</u>
lection	<u>lect1+ion]1</u>
lector	<u>lect1+or]2</u>
lecture	<u>lect1+ure]</u>
prelect	<u>[pre+lect1</u>

lect2 “leave out” (Total: 3), as in:

catalectic	<u>[cata+lect2+ic]1</u>
hypercatalectic	<u>[hyper+[cata+lect2+ic]1</u>

lect3 “entrap, snare” (Total: 7):

delectable	<u>[de+lect3+able]</u>
delectation	<u>[de+lect3+ation]</u>

lect4 “lick” (Total: 2):

electuary	<u>[e/x+lect4+uary]2</u>
-----------	--------------------------

---

<sup>5</sup> Calvert Watkins, *The American Heritage Dictionary of Indo-European Roots*, 3<sup>rd</sup> edn. Boston, New York: Houghton Mifflin, 2011.

<sup>6</sup> Julius Pokorny. *Indogermanisches Etymologisches Wörterbuch* (Bern, 1959), available at Leiden University’s website <http://www.indo-european.nl/> and at the University of Texas’ *Indo-European Lexicon* at <http://www.utexas.edu/cola/centers/lrc/ielex/PokornyMaster-X.html>.

lect5 “gather, choose” (Total: 146):	
analects	<u>[ana+lect5+s]3</u>
collect	<u>[co/m+l+lect5</u>
eclectic	<u>ec2+lect5+ic]1</u>
elect	<u>[e/x+lect5</u>
intellect	<u>[inte/r+l+lect5</u>
lectin	<u>lect5+in]03</u>
neglect	<u>neg+lect5</u>
predilection	<u>[pre+[di1+lect5+ion]1</u>
recollect	<u>[re+[co/m+l+lect5</u>
select	<u>[se+lect5</u>

(ii) As illustrated above, appealing to definitions tends to clarify differences among homographic longest strings, and appealing to etymologies can provide insights into how to cluster those homographs into groups.

The British philosopher John Austin maintained that words come to us trailing clouds of etymology.<sup>7</sup> And etymology provides extremely useful insights into elements and their explication, as when deciding on the boundaries between elements. Particularly useful is the *American Heritage Dictionary*, which has good etymologies and includes information on Indo-European roots and their descendants in modern English – and in the 5<sup>th</sup> edition, a similar list of Semitic roots. Although etymology does not dictate explication, it is prudent to contradict etymology as little as possible – much the same as with Occam’s Razor.

(iii) As discussed in *American English Spelling*, the Rule of syllabicity

urges that an element be at least a syllable in length. Such a rule is not as arbitrary as it may at first appear to be. Since the syllable plays an immediate and concrete role in spoken language, it seems only natural and efficient to expect the elements to follow the seams and boundaries set down by the syllables. Doing so lends to the element the concreteness and immediacy of the syllable. . . . This Rule of Syllabicity can be formulated even more stringently: The element tends to be no less and no more than a syllable long; indeed, the length of most elements is exactly one syllable.<sup>8</sup>

The sense of immediacy, or salience, can be experienced by comparing two words – one in which the syllables and elements are coterminous, the other in which elements are divided by syllable boundaries:

Elements are much more accessible [that is, salient] in words where their boundaries coincide with syllable boundaries, as compared with words whose elements cut across syllable boundaries and vice versa. In a word such as

---

<sup>7</sup> “A Plea for Excuses” in *Philosophical Papers*, 2<sup>nd</sup> edn. J. O. Urmson and G. J. Warnock, eds. London and Oxford: Oxford UP, 1970, pp. 201-02.

<sup>8</sup> D. W. Cummings. Baltimore and London: Johns Hopkins UP, 1988. p. 60. Hereafter cited as *AES*.

*restitute*, for instance, there is little parallelism between syllable and element, and the elements can be quite inaccessible: The syllabication of *restitute* is /res tə ˌt(y)ūt/; its explication into elements is [re+stit+ute]. Compare this with a word in which syllable and element boundaries coincide – for instance, *rebroadcast*: /rē ˈbrɒd ˌkɑst/, [re+broad+cast]. (AES 60)

Syllable boundaries often divide element boundaries and obscure the identity of elements. The spoken language can tend to obscure the structure of the written language – as with the colloquial *gonna* and the elements it merges in the words *going to*.

The stipulation that elements should be at least one syllable long is much stronger than the stipulation that they should not be longer than one syllable – as is shown in the list in the All Bases data table of the Lexis database. By adhering to the Rule of Syllabicity and thus working to keep elements of the written language coterminous with the syllables of the spoken language, we again strive for a type of simplicity and economy – as well as increased salience.

**(iv)** Another aspect of the quest for economy and simplicity is the need to minimize complex procedures. Since procedures have to do with the way elements and particles concatenate, they determine internal boundaries. The most common procedure is simple addition, by which elements concatenate with no changes – for instance, *repainted* [re+paint+ed]. Complex procedures – final <e> deletion (as in *delet<sup>e</sup>ion*), twinning (as in *twin+n+ing*), prefixal assimilation (as in [ad+s+simil+at<sup>e</sup>]1+ion]1) – all involve deletion, insertion, or replacement. So another part of the quest for simplicity is to minimize the number of complex procedures.

**(v)** In a written word a vestige is usually one or two letters that remain from an earlier stem – usually Latin or Greek – and that have lost their original semiotic, morphological, or syntactic content and function. In explications vestiges stand alone, with “+” signs immediately preceding and following them – empty holes in the middle of the word. An orthographic vestige is assimilated by merging it with a preceding or following element, to form a functioning expansion of the original element. To refer to this process, we take advantage of the polysemy of the word *assimilate*: “to become more similar,” as in the assimilation of final consonants in prefixes, and “to be incorporated or absorbed into,” as in the assimilation of immigrants into their new culture. The assimilation of vestiges leads to expanded bases of two or more syllables, which does violence to the weak form of the rule of syllabicity, but the payoff is fuller articulation in the explication.

In earlier versions of Lexis I tried to explicate to the smallest identifiable units, which unfolded far too many unassimilated vestiges. Their separateness led to a highly granular analysis – one with many small parts that often had little sense of connection among them – little articulation of parts. In the current version of Lexis I have tried to assimilate many of these vestiges by merging them with adjacent bases or suffixes. The purpose of explication is to unfold the most useful information as possible, but an unassimilated vestige carries only information about things in the distant past, no information useful to the present. But an assimilated vestige becomes part of a meaningful element, incorporated into the contemporary orthographic system. It is no longer rattling around meaninglessly in the word, no longer an empty hole in the middle of the word. It is more fully articulated with the other parts of the word, syntagmatically

and paradigmatically. It moves the explication of the word toward a more perfect articulation, a maximum simplicity.

The assimilation of vestiges can, I think, offer some insight into lexical evolution. Since nearly all orthographic vestiges are remnants of old stem forms, they are much like the vestigial organs that fascinated Darwin and are also like the so-called “junk” DNA in the genome, representing once-functional material that no longer serves its original purpose. Some modern particles are essentially vestiges that have evolved new functions – for instance, the particles *i1* and *o4* typically mark element boundaries, break up consonant clusters, and thus smooth the rhythm by breaking up the concatenation of clashing consonants. Further, the particle *o4* often takes the word stress, as in *speedometer*. Boundary marking, cluster breaking, stress bearing – these are newly evolved functions.

Orthographic vestiges between base and suffix often can be assimilated by being appended to the base. A good example of how far this process has already gone is the terminal <t> in Latinate bases like *fect*. Historically, the <t> is a vestige from the Latin past participle stem, but to explicate, say, *perfect* as \*[per+fec+t] would seem very forced – as well as contradicting the principles of syllabicity and economy. This is a more far-reaching issue than it at first may seem, for other examples of terminal <t>’s from Latin past participles occur in bases like *auct*, *caust*, *ceipt*, *ceit*, *cept*, *cerpt*, *cinct*, *debt*, *duct*, *fact*, *facult*, *fict*, *fact*, *flict*, *fluct*, *fract*, *frict*, *funct*, *gest*, *haust*, *ject*, *junct*, *pict*, *plot*, *punct*, *rect*, *rapt*, *sanct*, *script*, *sculpt*, *spect*, *stinct*, *struct*, *suct*, *tinct*, *tract*, *trait*, *unct*, *vect*.

There is much room for the further assimilation of vestiges. Those vestiges still in Lexis are one sign of the incompleteness of the process of explication. For more on vestiges see “Knowing what to do with vestiges,” item 2 in the “Open Questions” section below.

(vi) There is also a principle of productivity, based on whether an explication unfolds productive elements – that is, elements that occur in other words and thus lead to further paradigmatic relationships. Productivity is discussed in the first item in the “Open Questions” section below.

(vii) A prime desideratum is minimizing arbitrariness. An example of arbitrariness: We have the set {+phys1,+physi,+physio}. We also have the set {-a/1,-ia/}. So the placement of the <i> in the explication of a word like *hypophysial* seems quite arbitrary: [hypo+phys1+ial] or [hypo+physi+al]1. However, wider contextual considerations can sometimes help in such cases: The existence of the variant *hypophyseal* with -ea/ (together with the absence of an expansion <physe>) and the desire for consistency and parallelism urges the explication with +ia/, thus paralleling the -ea/ and reducing the arbitrariness.

However, there are many, many other cases of arbitrariness where so far there appears to be nothing in a wider context to urge one choice over the other. Surely there must be other helpful patterns. I think that the search for such patterns is an important problem for explication. Again, etymology should often help. Also helpful should be the desire to preserve semiotically close relationships, like that between the adjective and noun suffixes +otic] and +osis]. For instance, there are several elements that could be combined in various ways to create plausible explications of *narcotic* and *narcosis*: The

bases *+narc1*, *+narco1*, and *+narcot*; the suffixes *-ic]1* *-is]5*, *-os]5*, *-osis]*, *-ot]3* *-otic]*, *-sis]*, *-tic]*. This much choice produces many plausible combinations, and choosing which is the best leads to arbitrariness. However, since *-osis]* commonly contains the sense “condition, often diseased” in nouns (180 in Lexis) and *-otic]* is its companion adjective (107), there is strong pressure to choose *-osis]* and *-otic]*, in order to preserve the syntactic and semiotic parallelism – thus *narc+otic]* and *narc+osis]* with little or no arbitrariness.

Register may also be important: For instance, the technical-scientific register seems to be more inclined to expanded word-final elements -- such as *+ology*, *+imeter* and the like. I suspect this inclination is due to the fact that scientific words are created more or less on the fly, without the slow evolutionary process typical of ordinary language. (Brown's 'Foreword' and 'Introduction' to his *Composition of Scientific Words* contain a good discussion of how he thinks this on-the-fly composition should take place.<sup>9</sup>)

When faced with arbitrariness -- as, for instance, in the choice between *[pan+sperm+ia]1* and *[pan+spermi+a]2* – it seems better in general to preserve the free base and choose the first option, in keeping with the desire to reflect the plausible development of the word: It seems likely that the composer of *panspermia* started with the free base *sperm* in mind rather than the bound *spermi*. The same would hold, I guess, with the choice between *sperm+icide* and *spermi+cide*.

(viii) Concerning the desire for a plausible development of the word through a sequence of derivations: Consider *cooperate*, *cooperation*, *cooperative*: *[co+oper3+ate]1*, *[co+oper3+at/e]1+ion]1*, *[co+oper3+at/e]1+ive]1*. To go to *+ation]*, *+ative]* would obscure the development from *cooperate* to the other two. This is partially a sensitivity to etymology, but it is also an aspect of the search for the longest common string and full articulation within a cognate family.

The concern for a rational and plausible development of the word is nicely illustrated with the suffix sequence *+ic]1+al]1+ly]1*. In Lexis 434 words contain the full sequence – as in *canonically* *canon+ic]1+al]1+ly]1*, which grows from the series *canon*, *canonic*, *canonical*, *canonically*. But there are 661 words in Lexis that end <ically> but not with the sequence *+ic]1+al]1+ly]1*. For several of these words there are no shorter forms ending in *-al]1*. For instance, we have *hydraulic* and *hydraulically*, but no \**hydraulical*. (The *OED* does list *hydraulical*, but labels it obsolete and shows only three citations, stretching from 1664 to 1792.) There are also several words in Lexis that end <ically> but have no shorter form ending in *-ic]1*. For instance, we have *farfical* and *farfically*, but no \**farficc*. (Again, the *OED* lists *farficc*, but labels it obsolete and rare and lists only one citation, from 1763.)

The missing forms with final *-al]1* are due to two features of our lexicon: First, it is regular in English when adding *-ly]1* to a word ending in *-ic]1* to insert <al> even if there is no intermediate form ending in *-al]1*, as is the case with *hydraulic* and *hydraulically*. And second, though sometimes the two adjectives ending in *-ic]1* and in *-ic]1+al]1* have slightly different senses, such as *historic* and *historical*, usually they are synonymous, in which case the longer form tends to disappear. I do not explicate all words that end

---

<sup>9</sup> Roland W. Brown. *Composition of Scientific Words*. (Washington, D.C: Smithsonian Institution Press, 1956).

<ically> as +ic]1+a]1+ly]1. In order to accommodate those missing intermediate forms I use the expanded suffixes *-ical]* and *-ally]*, as in *farcical* farc+ical] and *hydraulically* hydr+aul+ic]1+ally]. Also, since dictionaries do not always agree whether certain forms exist, and since users are free to coin missing forms to fill emergent needs, I've decided to set the universe of inclusion at the Words table in Lexis itself – that is, if a form is not included in Words, for the sake of this analysis it does not exist.

Back-formations raise a special contention: English contains hundreds of back-formations: *AHD* lists 280; *W3* lists over 900. Explicating so as to show a plausible development of a word can flip a back-formation on its head. For instance, *AHD* shows *accrete* as a back-formation from the prior *accretion*, but I explicate *accretion* to [ad+c+cret/e+ion]1, just the opposite of the etymology. There are several examples of the way in Lexis back formations are explicated in contradiction of etymology, as in this small sample: *ablate*, *accrete*, *cerebrate*, *demarcate*, *donate*, *escalate*, *grovel*, *luminesce*, *peddle*, *scavenge*, *unit*.

As is apparent from the foregoing, these seven methods often contend with one another – in fact, nearly always. To rely on one of the methods while ignoring the others leads to an unstable minimum simplicity. Since our quest for perfect articulation means the quest for maximum simplicities, one of the burdens placed on explication is the apparently never-ending attempt to resolve these contentions.

## Open Questions.

English orthography is a changing, evolving system. Thus, any explication is to a certain extent tentative, a snapshot or freeze-frame. And although explication is usually straightforward, with the elements, particles, and internal boundaries quite clear, there can be arbitrariness as explication draws sharp lines in places where boundaries may be ambiguous, tentative, and fuzzy. This arbitrariness leads to some open questions, four of which that have pestered me the most are discussed below:

**1. Knowing when to stop, or how “deep” or “far back” to go in the analysis** – that is, deciding when you are looking at one element and when at two or more. It's a question of knowing where to draw the boundaries. This is a recurring problem, but the following four examples can give some insight into the nature of the problem and some proposed resolutions:

(i) I explicate *both* as a simplex both, though it has two distinct etymological components. The first, <bo>, evolves from Old English *bā* “both”, which led to an earlier word for “both,” *bo*, cited well into the 14<sup>th</sup> century. The second etymological component of *both*, <th> evolves from OE *þā*, plural of *þæt* “that,” and is related to modern *the*. So the question is, “Should *both* be explicated as a simplex, both, or as a complex, bo+th?” I explicate it as a simplex partially because of the principle of syllabicity, but also because the explication bo+th is unproductive – that is, it does not unfold, or produce, any useful elements. (The <th> in *both* is historically unrelated to the ordinal-forming suffix *-th]1*, as in *sixth*; the noun-from-verb suffix *-th]2*, as in *growth*; and the archaic 3<sup>rd</sup> person singular marker *-th]3*, as in *doth* and *saieth*.) The simplex *both* is an example of the evolutionary recycling of older materials and the way in which

echoes or fossils of older forms persist, serving new functions. It is an example of the orthographic version of biological speciation.

In *both* neither of the etymological echoes is productive beyond the word *both*, so I explicate it as a simplex. But in *erst*, as in *erstwhile* er2+st]5+while only one of the explicated elements, the superlative suffix +st]5, is productive, appearing in 52 words in Lexis: *first*, *most*, *worst*, etc. That seems to me to be enough to warrant the er2+st]5 explication, especially since -st]5 contracts the highly productive superlative suffix -est]1. This in spite of the productivity issue raised by the fact that +er2 occurs only in *erstwhile* and the archaic *erst* “earliest, erstwhile”.

(ii) Words like *decorate* could be explicated to *dec* plus a vestigial *or]4*. However, in all 29 words in Lexis the <dec> is followed by the <or>: *decor*, *decorate*, *decorum*, *decorous*, etc., which suggests that the explication should go only as deep as *decor*, overruling the weak form of the principle of syllabicity in the name of economy and the desire to minimize vestiges.

(iii) Modern *anthem* derives from Late Latin *antiphona*, roughly “the returned voice.” Since *antiphona* would itself explicate to [anti+phon+a]2, one could argue for *anthem* as [an+them], though it is not entirely clear where the <t> should go, the <t> and <h> having merged to a consonant digraph spelling the fricative /th/. Even though doing so again violates the weak form of the rule of syllabicity, it seems better to treat this as a simplex, anthem, on the grounds that, again, the more complex explication does not unfold any productive elements and the modifications from Late Latin are so extreme that, in evolutionary terms, the lexical version of speciation has again occurred, and we are dealing here with a new simplex element. Indeed, those modifications are so extreme that it is difficult to recognize the orthographic homology between *anthem* and its cognate, the modern *antiphon*.

(iv) Similar questions arise with the singleton eleven, from OE *endleofan*, which historically would be end+leofan “one left over (ten)” and with twelve, which in OE had already simplified to *twelf* but which derives from an assumed Germanic compound meaning “two left over (ten).”

*Hecatomb* derives from Greek *hekatōn* + *bovs* “a hundred oxen.” So etymologically it should be \*hecatom+b, but seems better left as a singleton, hecatomb.

The <tw> in *twelve*, is paradigmatic with the <tw>’s in simplexes like *two*, *twin*, *twine*, *twig*, all of which are from PIE *\*dwo* “two” and are instances of what I treat as subelemental patterning – patterning that cuts across different elements. Subelemental patterning is also found in several English adverbs and pronouns of location:

Where? (What place?)	Here (This place)	There (That place)
Whither? (To what place or direction?)	Hither (To this place)	Thither (To that place)
Whence? (From what source?)	Hence (From this source)	Thence (From that source)
When? (At what time?)	Now (At this time) <sup>10</sup>	Then (At that time)
What? (singular)	This	That
What? (plural)	These	Those

Applying the principle of the longest common string, these adverbs and pronouns could be explicated to include the ostensible elements *wh+*, *th+*, *h+*, *+ither*, *+ence*, *+ere*, *+en*, and less compellingly *+is*, *+at*, *+ese*, *+ose*. There are etymological cases to be made (i) for the *wh+*, which descends from the Indo-European *\*kʷo-*, the stem of relative and interrogative pronouns; (ii) for the *th+* from PIE *\*to*, the stem of demonstrative pronouns; and (iii) for the *h+* from PIE *\*ko*, another demonstrative pronoun stem. The main constraint against such explications is again the strong form of the rule of syllabicity. It seems better to treat these adverbs and pronouns as simplexes that contain instances of subelemental patterning, or the “striking homologies due to community of descent” mentioned by Darwin in *The Descent of Man* – in short, further examples of lexical speciation.<sup>11</sup>

Related to this problem of depth of analysis is the problem posed by groups like *temp1*, *tempor*, *temple2*, which in Lexis are treated as a single set. The underlying sense here is “time.” (Concerning *temple2* “area on the side of the head”: the theory is apparently that you can see the pulse in the temples, keeping time.) The base *temp1* and its expansion *tempor* occur in words in which the basic sense is something like “properly mixed, balanced,” which some scholars see as separate from the basic “time” sense. Watkins does not include the “time” words like *tempo*, *temporary*, *contemporary* in his list of words from PIE roots. The *OED* mentions the distinction but seems to favor collapsing the two into one, possibly by the progression “time” > “season” > “proper season” > “proper balance” (at *temper*). Partridge also collapses them, with the unifying notion of cutting or division into parts.<sup>12</sup> Lexis follows Partridge’s treatment.

Also related to the depth problem are groups of suffixes – especially polysyllabic suffixes – that are partners in the derivation of different parts of speech. For instance, *+esia*], *+etic*], as in the noun-adjective pairs *anaesthesia*, *anaesthetic*; *akinesia*, *akinetic*; *analgesia*, *analgetic*, etc. Or the noun-adjective pairs with *+esis*], *+etic*]: *genesis*, *genetic*; *hypothesis*, *hypothetical*; *prosthesis*, *prosthetic*; *synthesis*, *synthetic*. Or even the noun-noun pair *synaesthesia*, *synaesthesia*. Similarly there is the set *+osis*], *+oses*], *+otic*], with singular noun, plural noun, adjective trios like the following: *antibiosis*, *antiboses*, *antibiotic*; *neurosis*, *neuroses*, *neurotic*; *sclerosis*, *scleroses*,

---

<sup>10</sup> Obviously *now* breaks the pattern. Exactly why this is so is not clear. However, the *OED* records a now-obsolete form *hen* and its various forms that may have been part of the original pattern and that persisted into the 15<sup>th</sup> century. Our *now* also has a legitimate etymology going back to pre-English Germanic.

<sup>11</sup> *The Descent of Man, and Selection in Relation to Sex*, Chapter III, “Language.” Available online at [www.gutenberg.org/ebooks/23000](http://www.gutenberg.org/ebooks/23000).

<sup>12</sup> Eric Partridge, *Origins: A Short Etymological Dictionary of Modern English*, 2<sup>nd</sup> edn. NY: Macmillan, 1959, at “temper”, paragraphs 16-18.

*sclerotic*. Given a choice with such sets, it seems better not to divide the members into shorter suffixes. Better to keep the sets intact, to heighten the sense of relationship and articulation. (However, contradicting this notion, I explicate *synthetic* to [syn+thet1+ic]1 rather than \*[syn+th+etic], partially due to the rules of syllabicity and productivity, *thet1*, occurring in 25 words in Lexis.) I suspect this kind of pattern-protection – with all of its complications – will prove to be increasingly important in explication. Another common, more native set: +er[02, +est[1, +ly]1, +ness] poses no problems since all four are monosyllabic – as in *cuter, cutest, cutely, cuteness; weirder, weirdest, weirdly, weirdness*.

## 2. Knowing what to do with vestiges.

**N.B.** Of all aspects of explication the question of vestiges has caused me the most indecision, staring-off-into-space, and mind-changing. In proofing the explications in Lexis, I have tried to keep up with all the changes, in an effort to keep things as consistent as possible. But I'm sure I've missed some changes – probably many – leading to inconsistencies. For this, *mea culpa*.

If a nonterminative string conveys any contemporary information at all – semiotic, syntactic, morphological – I don't treat it as a vestige: Thus, -u]1 “marks diminutives” as in *formula* form1+u]1+a]2; or -yc] “nonterminative form of terminative -yx]”, which marks singular nouns from Greek – as in *calyces* and *calyx*, cal4+yc]3+es]3, cal4+yx]. Such forms are not vestiges since they convey information about current semantic and morphological relationships. I treat all vestiges as being empty and nonterminative, the idea being that if they are terminative, they assume some content, usually morphological or syntactic. For instance, the vestigial nonterminative suffix -ac]2 – as in *iotacism* iot1+ac]2+ism] and *portulaca* port2+u]1+ac]2+a]2 – contrasts with the terminative suffix -ac]4, which marks adjectives derived from Latin, as in *cardiac* cardi+ac]4 and *hypochondriac* [hypo+chondri+ac]4.

A good example of Occam's Razor at work involves the 32 words in Lexis with the base *er1* “of the Greek god of love Eros, sexual.” If in addition to the base *er1* we assimilate the vestigial -ot]3, thus producing the expanded base *erot*, we can explicate all 32 words with just the two base forms – *er1* and *erot* – with no unassimilated vestiges left over:

erogenous	<u>er1+ogen+ous</u> ]
pederast	<u>ped2+er1+ast</u> ]1
pederasts	<u>ped2+er1+ast</u> ]1 <u>+s</u> ]3
pederasty	<u>ped2+er1+ast</u> ]1 <u>+y</u> ]3
autoerotic	<u>auto+erot+ic</u> ]1
autoeroticism	<u>auto+erot+ic</u> ]1 <u>+ism</u> ]
autoeroticisms	<u>auto+erot+ic</u> ]1 <u>+ism</u> ]3 <u>+s</u> ]3
autoerotism	<u>auto+erot+ic</u> ]1 <u>+ism</u> ]
autoerotisms	<u>auto+erot+ic</u> ]1 <u>+ism</u> ]3 <u>+s</u> ]3
erotic	<u>erot+ic</u> ]1
erotica	<u>erot+ic</u> ]1 <u>+a</u> ]3
erotically	<u>erot+ic</u> ]1 <u>+ally</u> ]
eroticism	<u>erot+ic</u> ]1 <u>+ism</u> ]

eroticist	<u>erot+ic]1+ist]1</u>
erotization	<u>erot+ic]1+iz/e]+ation]</u>
eroticize	<u>erot+ic]1+ize]</u>
eroticized	<u>erot+ic]1+iz/e]+ed]1</u>
eroticizes	<u>erot+ic]1+iz/e]+es]1</u>
eroticizing	<u>erot+ic]1+iz/e]+ing]1</u>
erotism	<u>erot+ic]1+ism]</u>
erotisms	<u>erot+ic]1+ism]+s]3</u>
erotize	<u>erot+ic]1+ize]</u>
erotized	<u>erot+ic]1+iz/e]+ed]1</u>
erotizes	<u>erot+ic]1+iz/e]+es]1</u>
erotizing	<u>erot+ic]1+iz/e]+ing]1</u>
erotomania	<u>erot+ic]1+o4+man2+ia]1</u>
erotomanias	<u>erot+ic]1+o4+man2+ia]1+s]3</u>
homoerotic	<u>homo+erot+ic]1</u>
homoeroticism	<u>homo+erot+ic]1+ism]</u>
homoeroticisms	<u>homo+erot+ic]1+ism]+s]3</u>
homoerotism	<u>homo+erot+ic]1+ism]</u>
homoerotisms	<u>homo+erot+ic]1+ism]+s]3</u>

On the other hand, if we were to assimilate the vestigial <ot> to the common terminative suffix *-ic]1*, to form the expanded suffix *-otic]*, 16 of the 32 could be explicated to *er1* plus *-otic]*. This *-otic]* would be a homograph of the *-otic]* that forms Latin adjectives that partners with the noun-forming *-osis]* – as in *hypnosis*, *hypnotic*. *Erotic*, of course, is an adjective, but there is no <erosis> with the noun forming *-osis]*, the nearest noun form being *eroticism*. However, to explicate all 32 without any unassimilated *+ot]3* vestiges, in addition to *er1* and *-otic]* we would still need *erot*. In short, we don't need *-otic]* with these words. We can get by with only two forms – *er1* and *erot* – rather than the three with the expanded *-otic]*. Occam's Razor again.

Two other examples involve that vestigial *-ac]2*: (i) 11 words in Lexis contain the string <ostr> “bone, shell.” In all 11 the <ostr> is followed by <ac>, which could be treated as a vestigial *-ac]2*, but since every word with <ostr> has <ostrac>, it seems better to assimilate the <ac> to the <ostr>, giving us the expanded base *ostrac*. Doing so eliminates vestiges without increasing the number of bases. Similarly the string <psitt> “parrot” occurs in five Lexis words, all with <psittac>, leading to the base *psittac*, which like the expanded base *ostrac* is consistent with the Greek and Latin sources, in which the <ak> or <ac> is definitely part of the stem.

There can be indecision about whether to append a vestige to the preceding base or to the following suffix. One can argue that stronger-tends-to-attract, so, since bases are the semiotic cores of their words, the bases can be said to take precedence. If you think of a word as a massing of sound and sense around a central, attractive peak, this peak, or centroid, is always in ordinary usage the base. It is almost like a gravitational center that attracts loose or otherwise unattached material to itself – cosmologically speaking, like an early planetoid, gradually attracting free-floating material to itself to grow and increase in gravitational strength.

On the other hand, an argument for appending vestiges to following suffixes is that to do so continues a tradition of redivision common from the earliest times in Latin. Consider the following from the front matter of Glare's *Oxford Latin Dictionary*:

One of the most characteristic features of Latin suffixes is their growth by misdivision: for instance, the elementary suffix *-nus* gives rise to a group of secondary suffixes *-ānus*, *-īnus*, *-ernus*, *-tinus*. All such, if sufficiently common, are listed separately; where there may be doubt about the correct allocation, words are treated as containing the more specific, or longer, form of the suffix (e.g. *arcanus* is analysed as *ARCA* + *-ANUS*, even though historically a more correct analysis might be + *-NUS*. It is in many cases impossible to determine the historical facts, since in the classical period *-anus* was clearly felt as living suffix, whereas *-nus* was no longer employed in new formations. (“xxiii, “VI. Note on Suffixes”)

So the decision between assimilating vestiges to preceding bases or following suffixes presents us with another case of contending desiderata and a source of arbitrariness.

Lexis includes many expanded suffixes that contain assimilated vestiges. The following is intended to be a complete list:

-aginous]	farraginous, oleaginous, viraginous; 5 instances	Expands -inous]
-amen]	cyclamen, duramen, foramen; 9 instances	Expands -men]
-ament]	armament, fundamental, ligament, ornament; 87	Expands -ment]
-atic]	hieratic, problematic, sciatic; 127	Expands -ic]
-atile]	saxatile, versatile, volatile; 24	Expands -ile]1
-ation]	affirmation, allegation, vocation; 1338	Expands -ion]1
-atious]	disputatious, flirtatious, vexatious; 14	Expands -ous]
-atism]	conservatism, systematism; 10	Expands -ism]
-atist]	comparatist, privatist, systematist; 7	Expands -ist]1
-ative]	conative, consultative, hortative, imperative; 214	Expands -i've]
-ative]	aromatize, emblemize, privatize; 34	Expands -ize]
-ator]	conspirator, gubernatorial, penetrator; 105	Expands -or]
-atory]1	laboratory, lavatory, sudatory; 30	Expands -tory]1
-atory]2	modifier, perspiratory, revelatory; 67	Expands -tory]2
-ature]	curvature, musculature, tabature, vasculature; 14	Expands -ure]
-der]	enjoinder, rejoinder, remainder; 9	Expands -er]1
-ekin]	ramekin; 2	Expands -kin]
-ema]	edema, emphysema; 9	Expands -ma]
-emat]	edemata, empyemata, mathematics; 22	Expands -mat]
-erate]	recuperate, reverberate, tolerate; 43	Expands -ate]1
-erior]	interior, exterior, anterior, posterior; 28	Expands -ior]1
-esis]	catechesis, exegesis, pseudocycosis; 137	Expands -sis]
-ice/]	lenticel, pedicel; 3	Expands -ce/]
-icill]	penicillin, penicillium; 11	Expands -cill]
-ific]	classification, scientific, vilification; 242	Expands -fic]
-ify]	amplify, speechify, justify, mystify, testify; 606	Expands -fy]

-ikī]	cannikin, mannikin, minikin; 10	Expands -kin]
-imen]	regimen; 1	Expands -men]
-imentf]	impediment, rudiment, sentiment; 78	Expands -mentf]
-imony]	acrimony, alimony, sanctimony, testimony; 23	Expands -mony]
-inous]	cartilaginous, leguminous, mucilaginous; 20	Expands -ous]
-ison]2	caparison, garrison; 8	Expands -on]3
-itude]	altitude, certitude, platitude, pulchritude;	Expands -tude]
-nast]	slimnastics; 1	Expands -ast]
-nist]	tobacconist; 2	Expands -ist]1
-oneon]	bandoneon; 3	Expands -eon]
-osis]	halitosis, mitosis, sclerosis; 179	Expands -sis]
-roon]	octoroon, vinegarroon; 4	Expands -oon]1
-sis]	sepsis, arsis, peristalsis; 25	Expands -is]1
-tain]	chieftain; 4	Expands -ain]
-tane]	simultaneous; 4	Expands -ane]2
-terie]	bijouterie, marqueterie, papeterie; 6	Expands -erie]
-tery]	phalanstery, psaltery, serictery; 7	Expands -ery]1
-tic]	dramatic, majestic, styptic; 102	Expands -ic]1
-tin]	indigotin, zeatin; 3	Expands -in]3
-tious]	bumptious; 3	Expands -ious]
-tive]	forgetive; 1	Expands -l've]
-try]1	deviltry, marquetry, pantry; 7	Expands -ry]
-umentf]	document, monument, tegument; 37	Expands -mentf]

**3. Knowing how to handle rudimentary elements.** Rudiments are empty syntagms left when other more straightforward elements are explicated out of a word. All rudiments have very obscure, if any, content. Such rudiments are quite common in new technical terms. Some examples: +amite in *rolamite*, +ble4 in *bleomycin*, ex7 in *cephalexin*, tylos in *tylosin*, and vanc in *vancomycin*. But there are some old, well-established words that also contain rudiments: dand2 in *dandruff*, dog4 in *sundog*, gar8 in *garboil*, kibb in *kibbles*, old2 in *threshold*, and pad3 in *padlock*. In all these cases dictionaries describe the rudimentary element as of obscure and uncertain origin and sense.

Over time rudiments can acquire semiotic content. For instance, explicating the word *keelson* “an extra timber that strengthens the keel of a ship” to *keel* and the variant *kelson* to *kel2* leaves us with the rudimentary element *son3*, with the obscure and questionable sense of “swine, timber”: *keel+son3*, *kel2+son3*. The *OED* calls *son3* obscure of sense and source, and Partridge (at “keelson”) and Weekley (at “keel1”) suggest it could be a variation of a Scandinavian word with the sense “sill”<sup>13</sup>. More than a century after *keelson* (1611) and *kelson* (1611), came *stemson* (1769) and *sternson* (1846). *Stemson* “a timber that connects the stem keelson, and apron of a ship”; *sternson* “a brace that connects and braces the keelson and stem of a ship”: *stem1+son4*, *stern2+son4*. The creation of the two later words suggests that *son* had

---

<sup>13</sup> Ernest Weekley, *An Etymological Dictionary of Modern English*, vol. 1, NY: Dover, 1967 (Orig. edn. 1921).”

developed the sense “a supporting timber in a boat,” which would be a logical extension of one earlier sense of *sill* – though the connection with the questionable “swine” must remain obscure.

Rudiments can emerge due to pronunciation (and spelling) changes – for instance, *huckle* in *huckleberry*: Its original sense is unclear, though it appears to be an alteration of *hurtle* or *whortle*, as in *hurtleberry* and *whortleberry*, whose original senses are not at all clear-cut. The *huckle* in *huckleberry* appears to have no connection with *huckle*’s earlier senses of “hip, haunch,” “to haggle,” or “to bend.” *Huckle*’s current sense occurs only in the context of *huckleberry*, so that all it can really be said to contain is the marker that contrasts it with the first stems of other *berry* compounds: *rasp*, *straw*, *blue*, *goose*, etc. If someone were to create a new word (*huckletea*, for instance, for tea brewed from dried huckleberries), then *huckle* would contract the sense “huckleberry” and could be said, via synecdoche, to have become when used in that sense a true element, not a rudiment.

The label *rudiment* probably is also appropriate for elements in recently adopted words that have straightforward senses in their native languages but not yet in English. For instance, I explicate the Japanese word *ninja* to nin+ja, with the elemental Japanese sense “enduring” and “person,” but as yet little productive sense in English.

All in all, the issue of rudiments is much less complex and perplexing than the issue of vestiges and the next open question, the deletion of final <e>.

**4. Knowing when to explicate to silent final <e> deletions.** Next to vestiges I have had the most indecision and changes-of-mind when deciding when to explicate to a form with a silent final <e> that must be deleted in the word and when to explicate to its form without final <e>. For this current upgrade I’ve tried to follow the following principles concerning silent final <e> and its deletion:

When deciding whether to explicate to a terminative coform with a deleted silent final <e> or a nonterminative with no final <e>, the basic principle is this: If the word to be explicated either is or ends with a free stem that ends with silent final <e>, explicate to that stem, using final <e> deletion if necessary. Consider, for instance, the choice between terminative *jure* and its nonterminative form *jur*. The word *conjurer* contains the free stem *conjure* with silent final <e>, so its explication is [co/m+n+jur/e+er]<sup>01</sup> with the <e> deletion. Similarly:

<i>conjured</i>	<u>[co+n+jur/e+ed]</u>
<i>conjuring</i>	<u>[co+n+jur/e+ing]</u>
<i>conjunction</i>	<u>[co/m+n+jur/e+ation]</u>

And *conjures* [co+n+jure+s], with no deletion.

If the word to be explicated is not and does not end with a free stem that ends with a silent final <e>, use the nonterminative such as *jur*. Complicating matters here is the fact that the *OED* has entries for two nouns spelled <jure>, one of which has the sense “jurisprudence” and is chiefly Scots and obsolete, it’s most recent citation dated 1556. The other “a just privilege, a right” has only two quotations, one in 1533, the other

c.1745. *Jure* appears as a verb in the *OED* and in *W3* with the sense “to make of juror of” – but the *OED* gives only one citation, from Shakespeare’s *Henry IV, Pt. 1*: “You are grand iurers, are ye, weelee iure ye faith.” And *W3* labels *jure* a back-formation from *juror*. All in all it seems reasonable to conclude that there is no common free base *jure* that can be posited as the source of words like the following, so we use the nonterminative *jur*:

<i>jural</i>	<u>jur+al</u> 1
<i>jurist</i>	<u>jur+ist</u> 1
<i>juristical</i>	<u>jur+ist</u> 1+ic]1+al]1
<i>juror</i>	<u>jur+or</u> 2
<i>jury</i>	<u>jur+y</u> 3

A similar case involves the *lune, lun* “moon” set, which is somewhat more complicated than the *jure, jur* case. There are actually three words spelled <lune>, but none can be logically construed as the basis or source of the family of words: *Lune*<sub>1</sub> “leash for a hawk” is a variant of earlier *loyn* with the basic sense of “long” and nothing to do with the “moon” sense. *Lune*<sub>2</sub> “a fit of frenzy or lunacy” though related semiotically is late in English (*OED*’s first citation is from Shakespeare) and feels more like a derivative than a source. Much the same is true of the late and rare *lune*<sub>3</sub> “two arcs of a circle that enclose a space” and later “anything crescent-shaped.” The *lune, lun* nuclear family contains 39 words – *lune* and *lunes* plus 37 derivations. There is very little, if any, motivation to derive those 37 from *lune*. To do so would require a long string of ahistorical <e> deletions. It needlessly multiplies complex procedures with no etymological justification. Thus, all 37 are explicated to lun.

The *jure, jur* and *lune, lun* words illustrate the mistake of assuming that the shortest form with silent final <e> is always the more basic, the starting point. Actually silent final <e>’s – either usually or always – emerge in Middle English, rather late in the game. Often they are respellings of Old English inflectional suffixes that during Middle English weakened to a soft schwa sound. And several are late back-formations – as with the *fluve, fluv* family: The only word with *fluve* is *interfluve*, which is a 20<sup>th</sup> century back-formation from the 19<sup>th</sup> century *interfluvial*. In many cases the form with silent final <e> is a special case, reserved for terminative situations.

There is a subjective aspect: The principles outlined above hold only if the sense of the stem word is perceived as being the same as, or at least very close to, that of the derived word. If the form with silent final <e> is a free base, as it often is, then the choice between forms depends on semiotic relationships -- which of course introduces considerable subjectivity: If there is strong family resemblance between the free base and the affixed form, then explicate to the form with silent final <e>; if there is no or little felt similarity in senses, then explicate to the nonterminative form. The *mode, mod*<sub>4</sub> set offers examples of this complication: We have *modal* mod/e+al1], but *modest* mod4+est3], not \*mod/e+est], since the contemporary sense of *modest* is so far from that of *mode*. A word like *modern* seems to me to fall in between *modal* and *modest* in terms of family resemblance, with a sense like “in the contemporary mode”. With such tweeners I’m still inclined to explicate to the nonterminative form -- thus *modern* mod4+ern], avoiding a complex procedure. Obviously, derived forms would explicate to

the same coform as do the parent forms -- thus *modality* mod/e+al]+ity], *immodestly* [i/n+m+mod4+est]3+ly], *modernity* mod4+ern]1+ity].

**Some Conclusions.** There are many other problems and open questions in the explication of English words, but the foregoing are four of what I so far take to be the most important. Since orthography is a cultural-symbolic system, its evolution is Lamarckian, driven by imitation and learning. Explication can assist in that learning as it highlights unifying simplicities that can decrease a sense of arbitrariness and increase awareness of motivation within the system. Its work is complicated by the fact that our orthography is not only complex, but is also a living, evolving thing.

The work of explication is valuable, worthy of more attention than it has been given in the past in the general linguistic enterprise. Explication has a role to play as it explores those lexical interrelationships that drive symbolic reference, which Terrence Deacon argues is the distinguishing feature of human language. In Deacon's analysis symbolic reference, though it can, like simple reference, indicate some item in the world, does so not by simply pointing at it, but by drawing attention to it by way of the complex relationships between the word that names it and other words that make up the language. It is this interrelationship with other words that allows the referential power of a word to persist even in the absence of any matching of it with an external object of reference and thus distinguishes the word as symbol from the more primitive word as simple sign.<sup>14</sup> It underlies the difference between using the word *fire* symbolically in order to discuss fire in the abstract and using the word as a simple sign, as when you yell "Fire" in a crowded theater.

Explication can assist in the work of replacing a sense of arbitrariness in our language with a sense of the order and regularity – the motivation – that infuses it. An overbearing sense of arbitrariness is an enemy of learning. A strong assurance of the system's motivation is crucial if teachers and learners are to feel in control, to feel like agents rather than patients, like subjects rather than direct objects. We should remember Saussure's statement that "the whole system of language is based on the irrational principle of the arbitrariness of the sign, which would lead to the worst sort of complication if applied without restriction. But the mind contrives to introduce a principle of order and regularity into certain parts of the mass of signs, and this is the role of relative motivation" (133). And describing that order and regularity, that motivation, is a large part of the role of explication.

---

<sup>14</sup>*The Symbolic Species: The Co-evolution of Language and the Brain.* (NY and London: Norton, 1997) 43-46.