

Corpus-Based Research into Verb-Forming Suffixes in English: Its Empirical and Theoretical Consequences

Junya Morita

Kinjo Gakuin University/ 2-1723 Omori Moriyama-ku, Nagoya, Japan 463-8521
morita@kinjo-u.ac.jp

Abstract

The present study explores the semantic and structural aspects of word formation processes in English, focusing on how verbs are derived by the suffixes *-ize*, *-ify*, *-en*, and *-ate*. Based on relevant derivatives extracted from the British National Corpus, their detailed observation is made from semantic and formal viewpoints. Then their theoretical analysis is carried out in the framework of generative theory. The BNC survey demonstrates that (i) the meanings of derived verbs are largely divided into five types and the submeanings are closely related to each other, (ii) the well-formedness of derived verbs is primarily determined by the semantic and formal features of their bases, and (iii) *-ize* suffixation is creative enough to provide a constant supply for new labels. To account for these empirical observations, the mechanism for forming *-ize* derivatives is proposed in which the semantic properties and creativity of *-ize* derivation stem solely from the underlying structure and the formal properties of the bases derive from the lexical entry of *-ize*.

Keywords: corpus research, derived verbs, semantics, morphosyntax, word formation mechanism

1 Introduction

The central issue of generative morphology is how to account for children's lexical acquisition: they acquire the vocabulary rapidly and accurately based on limited and often degenerated data. The most promising way of achieving this is to

establish a general mechanism to generate an infinite number of possible words so that words to remember are greatly reduced in number. In addition, the mechanism itself needs to be of wide application and elegant in linguistic terms to minimize information specified in the grammar.

As part of the above enterprise, the present study attempts to construct a system which automatically produce well-formed derived verbs, as in “she has been hopelessly *sentimentalized* and hopelessly *magicalized* by tradition (BNC: ACL 1522).” This study is organized as follows: after outlining the method of research (section 2.1), we closely examine the derivation of verbs in English and illuminate its semantic features (sections 2.2 and 2.3). Then we elucidate its morphological properties—the formal restrictions of the bases and productivity (sections 2.4-2.6). Section 3 presents theoretical implications for the results of our research. A summary of the main arguments is presented in section 4.

2 Observation and Generalizations

2.1 Target and Methodology

In this section, we will make systematic observations of English derived verbs and present generalizations based on it. We now begin a brief description of the method of research and the resulting data. By repeatedly using the “wild card” function of a research engine, words ending in *-ize*, *-ify*, *-en*, and *-ate* are extracted from the British National Corpus (BNC), a 100-million-word corpus.¹ In particular, their frequency is checked to identify the hapax legomena (token frequency 1). As a result of the research, we have obtained 381 word types in *-ize*, 68 word types in

¹I am indebted to the research engine of www.english-corpora.org (BNC).

-ify, 58 word types in -en, and 447 word types in -ate, including 123, 5, 2, and 26 hapaxes, respectively.

2.2 The Meanings of Derived Verbs

This section and the next deal with the semantic aspects of verb derivation. The semantic aspects of derived verbs have been well observed in the literature from a descriptive perspective (Jespersen, 1949; Marchand, 1969) and from a generative perspective (Plag, 1999; Lieber, 2005). According to Plag (1999: 125), the meanings of derived verbs can be divided into seven classes: 1 resultative ‘make into x’; 2 locative ‘put (in)to x’; 3 ornative ‘provide with x’; 4 performative ‘perform x’; 5 similitive ‘act like x’; 6 causative ‘make x’; 7 inchoative ‘become x.’

We will revise this classification in the following way. First, ‘resultative’ and ‘locative’ may be classed together as ‘result’; *atomize* denotes ‘put sth in a state of an atom’ and *hospitalize* signifies ‘put sb into a hospital,’ and thus both of them are associated with the change of state or place.

Second, two more submeanings join the classification, ‘agentive’ and ‘instrumental’; *patronize* and *cauterize* are interpreted as ‘act as patron’ and ‘do sth with cauter,’ respectively. Then we may group ‘ornative,’ ‘performative,’

‘agentive,’ and ‘instrumental’ under the heading of ‘providing or giving.’ This is because they are all interpretable as ‘make sb/sth provided with x; *chemicalize* (ornative) means ‘give chemical to sb/sth,’ *dichotomize policemen* (performative) signifies ‘give or apply the process of dichotomy to policemen,’ *patronize the shop* (agentive) represents ‘(in a widened sense) assign a patron to the shop,’ and *cauterize* (instrumental) denotes ‘provide sb with cauter.’ Finally, there is one other addition to the traditional classification; the submeaning “function,” referring to ‘make sth be as x,’ should be set up to interpret *canonize the texts* as ‘make the texts be as canon.’²

Table 1 shows the results of our research into the semantics of derived verbs.³ Deadjectival derived verbs are essentially transitive verbs and have the meaning of ‘make sth x’ (causative), as in *circularize* ‘make sth circular.’ We see in Table 1 that the submeaning ‘causative’ is the highest in number of word types and hapaxes, showing that this is the central meaning of derived verbs. Part of these derivatives can be used as intransitive verbs and they mean ‘become x’ (inchoative). This shift has been well studied from a generative perspective; it is frequently treated as an alternation between transitives and inchoatives (Levin and Rappaport Hovav, 1995). We will not, though, deal with the issue of how they are related to each other.⁴

| base | the meaning of derivative | -ize | -ify | -en | -ate | total |
|----------------|---------------------------------|-----------|--------|--------|----------|-----------|
| Adj | causative | 215 (67) | 22 (3) | 51 (2) | 19 (0) | 307 (72) |
| N | (a) result | | | | | |
| | (i) resultative | 51 (13) | 17 (2) | 4 (0) | 7 (0) | 79 (15) |
| | (ii) locative | 3 (0) | 0 | 0 | 0 | 3 (0) |
| (b) providing | (i) ornative | 35 (16) | 3 (0) | 1 (0) | 8 (1) | 47 (17) |
| | (ii) performative | 35 (8) | 3 (0) | 1 (0) | 4 (1) | 43 (9) |
| | (iii) agentive | 7 (4) | 0 | 0 | 2 (0) | 9 (4) |
| | (iv) instrumental | 5 (2) | 0 | 0 | 0 | 5 (2) |
| (c) similitive | | 15 (9) | 1 (0) | 0 | 0 | 16 (9) |
| (d) function | | 3 (2) | 1 (0) | 0 | 1 (1) | 5 (3) |
| | purpose | 1 (1) | 0 | 0 | 0 | 1 (1) |
| | bound stems | 11 (1) | 21 (0) | 1 (0) | 406 (23) | 439 (24) |
| | total number of types (hapaxes) | 381 (123) | 68 (5) | 58 (2) | 447 (26) | 954 (156) |

Table 1: The submeanings of -ize, -ify, -en, and -ate words

²We can find the submeaning ‘purpose’ (e.g. *winterize*), but this is quite exceptional.

³If a derived verb can be interpreted in two ways (e.g. *capitalize* ‘resultative/ornative’), it is separately counted. Cases of this kind are very few indeed—only 2 cases. Additionally, when the base can be an adjective or noun, the one which is naturally interpretable is chosen. For example, *editorialize* is denominal, since it means ‘to express an opinion in an editorial.’

⁴Levin and Rappaport Hovav (1995: 102-104) comment as follows: transitive verbs implying the intervention of an intentional agent do not have intransitive forms, as exemplified in (i), and -ize and -ify verbs are generally of this kind so that they cannot normally detransitivize, as illustrated in (ii).

- (i) a. The terrorist *assassinated* the senator.
- b. *The senator *assassinated*.
- (ii) a. The farmer *homogenized* the milk.

As for denominal derived verbs, the submeanings of ‘resultative,’ ‘ornative,’ and ‘performative’ form a majority in number of word types and hapaxes, and so they are the central meanings of denominal derived verbs. The productivity of each derived verb will be discussed in section 2.6.⁵

2.3 Correlation between Derived Verbs and Their Bases

This section shows how, and to what extent, derived verbs’ meanings are predicted from their bases. This issue has received relatively little attention in previous morphological analyses. We have already stated that deadjectival verbs generally have the causative sense ‘make sth x’ and part of them may become inchoatives. The meanings of denominal verbs are largely divided into four classes and each class is closely related to the meanings of the base. Importantly, the three main subclasses of meanings—‘resultative,’ ‘ornative,’ and ‘performative’—are generally predictable from the bases’ meanings. The meaning correspondence is offered in Table 2.⁶

‘resultative’

- (a) state/quality: *dimerize, fossilize, harmonize, isomerise, mylonitize, pauperize*, (19)
- (b) language: *capitalize₁, diarize, editorialize, journalize, novelize, poetize, rhapsodize* (10)
- (c) basic element: *atomize, carbonize, oxidize, unitize* (4)
- (d) system/classification: *categorize, factorize, computerize, memorize, organize* (21)
- (e) one’s status: *deputize₁, idolize* (2)

‘ornative’

- (a) (bio)chemical substance: *chemicalize, heparinize, narcotize, siliconize, trypsinize* (8)
- (b) presentable thing: *accessorize, capitalize₂, caramelize, deodorize, subsidize* (12)
- (c) academic matter: *anthropologize, biologize, botanize, philosophize, symbologize* (7)
- (d) format: *alphabetize, rasterize, tokenize* (3)
- (e) rights: *autonomize, hegemonize, prioritize* (3)

‘performative’

- action/process: *anatomize, apologize, eulogize, assassinize, dichotomize, economize* (32)

Table 2: Meaning correlation between derived verbs and their bases

Three points are worth noting here. First, a derived verb denotes ‘resultative’ when the base noun has one of the five meanings; if an underlying noun (*dimer*) expresses a state or quality, the derivative (*dimerize*) can naturally be taken as meaning ‘put sth in a state/quality.’ From nouns related to language are derived verbs that denote ‘put sth into a verbal form,’ as in *diarize*. Nouns indicating basic elements and those conveying system/classification are verbalized to mean ‘put sth into a basic element’ (*atomize*) and ‘put sth into a system/classification’ (*categorize*), respectively.

Second, the given meanings of base nouns lead to the meaning ‘ornative’ of the derivatives; from the names of (bio)chemical substance are derived verbs that signify ‘give the substance,’ as in *chemicalize*. This correlation is reasonable, since chemical substance is usually given to somebody or something to cause chemical action. Examples such as *accessorize* and *capitalize₂* ‘provide (a company) with capital’ can be treated similarly; accessory and capital are presentable, that is, suitable to be presented. In addition, the ‘ornative’ meanings of derived verbs are commonly expected from the underlying nouns indicating academic matter, format, or rights. For instance, *anthropologize* and *alphabetize* imply ‘provide an anthropologic view’ and ‘provide an alphabet format,’ respectively. Finally, we can easily assign the ‘performative’ meaning to a derived verb when the base noun entails an action or process, as in *anatomize*.

Turning to other submeanings, we can easily understand that a verbal suffix combines with a noun expressing a place to produce a verb with a sense of locative (*palletize*) and a verbalizer is added to an agentive noun to form a verb that refers to the related action (*burglarize*). From nouns of instruments are derived verbs that denote the action for which the instruments are meant (*catheterize*) and ‘similative’ verbs are built from proper nouns (*Beethovenize*) and animal names (*serpentize*).

b. *The milk *homogenized*.

⁵As evidenced in *refer, remit, and resume*, bound stems generally have no fixed meanings; only words may have constant meanings. Thus, “all regular word-formation processes are word-based” (Aronoff, 1976: 21). According

to this thesis, a case in which a verbal suffix attaches to a stem is left out of consideration here.

⁶The number in a parenthesis indicates the total number of word types.

2.4 Formal Restrictions on the Bases

The previous sections have examined semantic facets of verbalization. In this section and the following two sections, we will demonstrate the morphological facts on derived verbs. We will concentrate here on the internal structure of the bases and their vocabulary strata. Restrictions are imposed on the size and composition of the bases. First, as is pointed out in Marchand (1969: 100), verbal suffixes do not combine with compounds and this is attributed to the general inhibition of direct verb compounding (cf. **to rock-throw*). Our research supports this view; there are no such verbs in BNC (**rock-crystalize*, **rock-solidify*, **knife-sharpen*). Second, as a result of the same research, we find that a verb-forming suffix generally does not attach to prefixed bases. Thus, *-ize*, *-ify*, *-en*, and *-ate* do not combine with words including prefixes such as *a-*, *trans-*, and *ultra-* (**atypicalize*, **transcontinentalize*, **ultratrendify*). We have only three counterexamples to this: *immobilize*, *impersonalize*, and *internationalize*.

Third, Lieber (2005: 412) states that the verbalizers *-ize* and *-ify* normally do not attach to suffixed words, excepting those ending in *-al*, *-ian*, and *-ic*. However, our research demonstrates that there is considerable variation in the combination of suffixed words between verbal suffixes. *-Ize* attaches to words ending in *-able* (*permeabilize*), *-ive* (*passivize*), *-er* (*computerize*), *-(a)(t)ion* (*revolutionize*) in addition to *-al*, *-(i)an*, and *-ic* bases (*commercialize*, *Christianize*, *classicize*). Some suffixes in verbal bases are truncated when combined with *-ize*, as exemplified by *-ous* in *anonymize*. Morpheme truncation will be discussed in the next section. Contrastively, other verbal suffixes can attach to suffixed bases in a very limited way. *-Ate* can attach to *-al*, *-ant*, *-ic* and *-ous* bases (*liberate*, *resonate*, *rubricate*, *stimulate*), *-ify* can be added to *-ic* and *-ity* bases (*mystify*, *commodify*), and *-en* can affix to *-(i)an* and *-th* bases (*Christen*, *strengthen*). Most of these base-internal suffixes are truncated in combination with the verbal suffixes. We can say then that *-ize* affixation is a major verb-forming process in the sense that it may attach to various suffixed bases to produce a variety of verbs.

Let us now turn to the issue of vocabulary strata. It has been well observed that an affix chooses an

item of a specific vocabulary stratum; *-ize*, *-ify*, and *-ate* typically combine with words of Latinate origin, while *-en* normally combines with words of native origin (Jespersen, 1949; Marchand, 1969). Our BNC research has identified the vocabulary strata of words with which each suffix combines: (i) [Latinate] (354 word types), [Greek] (18), [Native] (7), the others (2) for *-ize*; (ii) [Latinate] (64), [Native] (3), the others (1) for *-ify*; (iii) [Native] (53), [Latinate] (5) for *-en*; (iv) [Latinate] (447) for *-ate*. The result leads us to conclude that *-ize* mostly takes [Latinate] or [Greek] bases, *-ify* and *-ate* predominantly or exclusively take [Latinate] bases, while *-en* mainly takes [Native] bases. Thus, the previous observations have been confirmed by our BNC research.

It is widely accepted that affixes can be divided into two classes: one may cause phonological change of the base (class I), while the other is phonologically neutral (class II). Additionally, their ordering is recognized: class I affixes cannot appear outside class II affixes. *-Ize* may be considered as a class I affix, since it may change the phonological quality of the base (cf. *stable* and *stabilize*). According to Selkirk (1982: 81), the suffixes *-ful*, *-less*, *-ly*, *-y*, *-ish*, *-en*, *-ed*, *-some*, *-able*, *-er* are all class II suffixes, and hence they are predicted not to occur in *-ize* derivatives. This prediction is confirmed by the ill-formedness of words such as **harmfulize*, **powerlessize*, and **friendlyze*, which are never found in BNC. It is worth noting here that all the suffixes except *-able* are of native origin. The co-occurrence restriction is then deduced from the requirement that a base be largely Latinate or Greek, and therefore the present ordering will be unnecessary for *-ize* verbalization.⁷

2.5 Truncation of a Word-Internal Suffix

This section deals with the truncation of a word-internal suffix concerning verbalization, focusing on *-ize* affixation. There are good reasons for the truncation of an intra-word suffix. One is that the underlying form of *[X-suffix]-ize* is well suited to the meaning of the whole word. For example, *systematize* means ‘make sth systematic’ and so the meaning is assigned easily and naturally to the word if *-ic* is underlyingly involved in the word. Another is that we can get rid of an unnecessary bound base; the lexicon would be redundant if the

⁷Selkirk (1982: 81) points out that *-able* has dual status, that is, it may be a member of both classes. It might be argued

then that the type of the affix *-able* involved in *-ize* derivatives belongs to class I.

bound stem *systemat-* were listed only for *-ize* affixation. The strongest reason of all is the fact that there exists a doublet of truncated form and untruncated form, as exemplified in *digitize/digitalise* and *monetize/monetarize*. There seems to be no significant meaning difference between both forms, and hence their relationship can be described clearly by the relevant truncation.

With respect to suffix-containing *-ize* verbs, some internal suffixes are truncated while others are not. The results of our BNC survey are shown in Table 3.

truncated suffixes: *-ic* (20 types), *-ous* (4), *-al* (3), *-ity* (3), *-ant* (1), *-ism* (1), *-ive* (1)
untruncated suffixes: *-al* (81), *-ic* (13), *-(i)an* (8), *-able* (3), *-(a)(t)ion* (2), *-ary* (1), *-er* (1)

Table 3: Truncation of a suffix in *-ize* words

Seven suffixes are deleted in *-ize* verbs: *-ic* (e.g. *anaesthetize*), *-ous* (*anonymize*), *-al* (*attitudinize*) *-ity* (*authorize*), *-ant* (*deodorize*), *-ism* (*ostracize*), and *-ive* (*sensitize*). By contrast, seven suffixes prove to be intact in *-ize* verbs: *-al* (e.g. *centralize*), *-ic* (*classicize*), *-(i)an* (*Americanize*), *-able* (*respectabilize*), *-(a)(t)ion* (*productionize*), *-ary* (*militarize*), and *-er* (*computerize*).⁸

-Ic truncation deserves special mention. *Ic-* is essentially deletable in the position at issue; twenty word types of such derivatives are identified in BNC. However, we detect thirteen word types of derivatives whose internal *-ic* is not deleted: (i) *classicize*, *ethicize*, *Gallicize*, *Gothicize*, *poeticize*, *publicize*, (ii) *romanticize*, *geometricize*, (iii) *aestheticize*, *cosmeticize*, *eroticize*, *hermeticize*, *phonemicize*. Looking closely at these examples, we notice that the base of the internal suffix *-ic* is monosyllabic as in (i) and it ends in two consonants as in (ii). Then, a generalization emerges: when the base of the internal suffix *-ic* is polysyllabic or ends in a single consonant, *-ic* truncation applies. Although the examples in (iii) remain unaccounted for, the generalization applies to the *-ive* truncation as well (cf. *passivize* and **passize*).

To conclude this section, the internal suffix *-al* is generally intact in *-ize* derivatives while suffixes like *-ous* and *-ity* are truncated. The suffix *-ic* may

be either truncated or untruncated and a generalization can be made about the truncation process at work.

2.6 Productivity

As the last morphological facet, we will discuss the productivity of verb-forming suffixes. A hapax-centered productivity measure for derivation is applied to data collections to calculate the productivity value of verb-forming process. We accept a hapax-based productivity measure, which gives a key role to hapax legomena of a large-scale corpus (Baayen and Renouf, 1996). This rests on the view that the capacity of an affix to create new forms crucially involves the degree to which the affix yields words of ultra-low frequency (Hay, 2003).

We propose a productivity measure: $Productivity (P) = n_i/V$, where n_i is the number of hapaxes and V is the total number of word types.⁹ Our BNC research detects 123 hapaxes and 381 word types of *-ize* derivatives, giving its productivity value of 0.323 (cf. Table 1). In this measure, the productivity of *-ize* affixation is defined as the potentiality of creating 123 kinds of new words when 381 kinds of *-ize* derivatives are used; nearly one-third of the attested *-ize* types are innovated verbs. According to the same measure, the productivity values of *-ify*, *-en*, and *-ate* verbalization are, respectively, 0.074, 0.035, and 0.058. The results of the research then demonstrate that while *-ify*, *-en*, and *-ate* are not productive affixes, *-ize* affixation is fairly productive to promote the creation of neologisms.

Additionally, *-ize* derivatives may be created depending on context. In example (1), the process of making worms into arthropods is momentarily lexicalized with the verb *arthropodize*, relying on the preceding noun *arthropods*. Example (2) illustrates how a complex word is created in the enumerative or listing environment; a series of comparable activities are enumerated by the use of three *-ize* final verbs, with *moronised* and *lobotomised* being innovated. Online word formation at issue is largely determined by the functions of “naming” (to conceptualize a property by giving it a name) and “brevity” (to construct a concise and sensible word) (Clark and Clark, 1979;

⁸We confine our attention to well-established and recognized suffixes, that is, those listed in Quirk et al. (1985: 1548-1555). Hence we leave out of consideration suffixes like *familiar*, *alkaline*, and *maximum*.

⁹This productivity measure is a revised version of the one proposed by Baayen and Renouf (1996), who place the total number of tokens in the denominator of the productivity formula.

Rice and Prideaux, 1991). That *-ize* words may be constructed wherever there exist such functional requirements confirms the derivational potentials of the verbalizer investigated.

- (1) ... different *arthropods* may have come from different and separate worms, independently, which became “*arthropodized*” by acquiring an external skeleton. (BNC AMM: 953)
- (2) She describes women, for example, as “*moronised*,” “*robotised*,” “*lobotomised*,” as “the puppets of Papa.” (BNC ECV: 1405)

3 Theoretical Perspectives

3.1 Antilexical Approach

Our task in this section is to formalize *-ize* affixation, a major verb-forming process. Specifically, we will present pertinent syntactic structures, lexical entries, and subsidiary rules. Before proposing a new analysis, let us sketch a grammatical model on which our analysis relies.

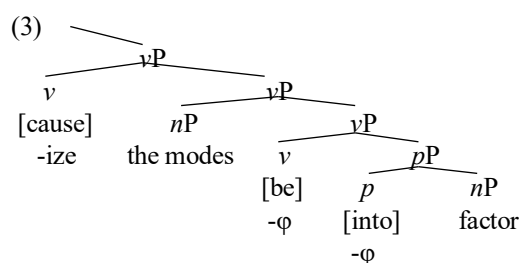
The properties of *-ize* derivatives observed above are best dealt with in the framework of antilexicalism. This thesis holds that major word formation processes take place outside the lexicon so that the creative aspects of sentence and word construction are uniformly captured in syntax (Halle and Marantz, 1994; Matushansky and Marantz, 2013). The creativity of *-ize* derivation substantiates the view that it is treated in syntax, but not in the lexicon, which is generally defined as a set of listed items. Thus, such a view has the merit of maintaining the homogeneity of the modules.

The present thesis also has the theoretical advantage of removing the *-ize* word formation rule from the lexicon by the independently established syntactic devices, whereby the related redundancy is expelled from the grammar completely. Moreover, an empirical advantage can be gained by adopting the antilexical approach. As indicated in section 2.6, the syntactic environments, anaphoric and enumerative, can be a major source of word creation. The spur-of-the-moment word composition in the syntactic contexts provides a constant supply for new labels like *arthropodize* and *moronise* and thus backs up the thesis of syntactic word formation.

3.2 Underlying Structures

Let us consider the underlying structures concerning *-ize* words in the framework of antilexicalism. We follow Hale and Keyser’s view that the meaning of a complex word is primarily determined by the syntactic structure (Hale and Keyser, 1993). Thus, the converted verb *to shelve* is arguably derived by the head-movement of a noun (*shelf*) from an underlying structure such as $[_{vP} -\phi [_{vP} \text{the book } [_{vP} -\phi(=\text{put}) [_{pP} -\phi(=\text{on}) \text{shelf}]]]]$, with abstract *v* and *p*. According to this view, the deadjectival *-ize* construction has the basic structure of $[_{vP} v_{[\text{cause}]-\text{ize}} [_{vP} [_{nP} \text{the cell}] [_{vP} v_{[\text{be}]-\phi} [_{pP} \text{alkaline}]]]]$, where the underlying adjective *alkaline* is a predicative of the surface object *the cell* and the “small clause” is dominated by the causative *-ize*. Thus, the causative meaning of *they alkalize the cell* can be readily obtained from the underlying configuration.

As observed in section 2.2, the meanings of denominal verbs are broadly divided into four types: (i) ‘make sth be into x’ (result, e.g. *factorize*), (ii) ‘make sth provided with x’ (providing/giving, *chemicalize*), (iii) ‘make sth be like x’ (simulative, *Beethovenize*), and (iv) ‘make sth be as x’ (function, *canonize*). Accordingly, the structure of the result-type will be as shown in (3):



The sentence *they factorize the modes* is then to be interpreted as meaning that they make the modes broken down into factors. The structure of the providing-type is essentially the same as that of the result-type: $[_{vP} v_{[\text{cause}]-\text{ize}} [_{vP} [_{nP} \text{the dress}] [_{vP} v_{[\text{be}]-\phi} [_{pP} p_{[\text{with}]-\phi} [_{nP} \text{accessory}]]]]]$. The only difference between the two types is that the providing-type involves the preposition *with* as opposed to *into*, so that the meaning of *they accessorize the dress* is something like ‘they make the dress accompanied by an accessory.’ Note that one of the main meanings of *with* is ‘accompanied by another person or thing.’

Similar remarks apply to the simulative-type and function-type of *-ize* derivatives. The former example *he Beethovenized Haydn’s minuet* has a

syntactic structure: $[_{VP} v_{[cause]-ize} [_{VP}[_{NP} \text{Haydn's minuet}][[_{VP} v_{[be]-\phi}[_{PP} p_{[like]-\phi} [_{NP} \text{Beethoven}]]]]]]]$, where the preposition *like* makes a difference in the way the base noun is characterized. From this follows the meaning: ‘he made Haydn’s minute like a work of Beethoven.’ To take the latter example, *they canonized the texts* has a configuration: $[_{VP} v_{[cause]-ize} [_{VP}[_{NP} \text{the texts}][[_{VP} v_{[be]-\phi}[_{PP} p_{[as]-\phi} [_{NP} \text{canon}]]]]]]]$. This type differs from others in that the preposition *as* is involved, so that the reading ‘they made the texts as a canon’ can readily be obtained.

There are two advantages of the present analysis. First, it can account for the meaning properties of derived verbs observed in section 2.2; the basic meaning and additional meanings of *-ize* verbs can be distinguished accurately. The basic one is ‘make y (be) in the state of x’ and this meaning is attributed to the core part of the *-ize* construction. The additional meanings are divided into five types according to what condition the surface object y is in. This is typically represented by the spatial and functional relations that are expressed by specific prepositions. Thus, the difference between the submeanings originates in the different prepositions in the core layer, whereby the submeanings can be related to each other.

The second advantage is that possible classes of *-ize* verbs can be predicted from our analysis: *-ize* verbs can only be transitives and ergative intransitives (inchoatives). Two cases in point can be recognized: unergative (intransitive) verbs do not engage in *-ize* affixation, as in **they dancize to rap music/I must journe(y)ize there*. This is because unergatives typically signify movement of animate entities and such a movement/action construction is not fitted to the predicative nature that *-ize* affixation involves. Note that converted verb may be a verb of this type (*they dance to rap music/I must journey there*), since verbal conversion does not necessarily involve predicative construction. Additionally, *-ize* derivatives of unaccusatives are illicit, as in **ethical problems will surfacize (=ethical problems will rise to the surface)/*lower level of pollution will resultize*. An unaccusative (intransitive) verb expresses a phenomenon that happens spontaneously without the intervention of any causer, which is incompatible with the intentionality that *-ize* verbs imply.

3.3 Vocabulary Insertion

Derived words are constructed by inserting an affix in an appropriate syntactic node based on its formalized lexical entries (Harley and Noyer, 2000; Embick, 2010). From the semantic and morphological properties identified in section 2, we can describe the internal features and selectional conditions of *-ize*: all five types of the suffix *-ize* have a common feature as verbalizer, yet each requires the base with a distinct feature. These descriptions can be formalized into the lexical entry on the basis of an underspecified model, as seen in (4).

- (4) *-ize*: (a) $[V][[cause]$, (b) $+< vP_{[be]}$,
 $a/p_{[into]}/p_{[with]}/p_{[like]}/p_{[as]}$, Latinate/Greek>
 Condition: predicative=[root (suf)]

The internal features of the affix are listed in (a) and its license environment is specified in (b). We here assume “Generalized subcategorization,” which enables subcategorization features to include not only the features of the whole category but also those of its lexical head and complement (cf. Emonds, 2000). The lexical entry *-ize* in (4) then designates something like ‘*-ize* makes a causative verb, adjoining to a “small clause” consisting of a subject and a predicative; the predicative is divided into five groups and they are all of Latinate or Greek origin.’ For instance, when *-ize* is inserted under the *v* node in the environment of predicative including $p_{[into]}$, the result-class of *-ize* derivative is obtained. The condition of predicative entails that compounds and prefixed words are ruled out as the base of *-ize*. The crucial point is that *-ize* verbs are freely coined as long as the affixation meets the licensing conditions, particularly those on the structure of the bases and their vocabulary strata.¹⁰

3.4 Subsidiary Rules

This section focuses on two kinds of auxiliary rules for *-ize* derivation. The first one is a “redundancy rule,” which eliminates the redundancy of item-by-item specification. As shown in section 2.3, there is an essential meaning correlation between *-ize* verbs and their bases. Confining discussion below to the resultative-type and ornative-type, we

¹⁰How to construct a word form from the corresponding syntactic representation will not be explored here. There are

two ways in which such a word is constructed: one is to use syntactic head-movement (Harley, 2009); the other is to use morphological merger (Marantz, 1996).

observe that *-ize* verbs with a sense of ‘resultative’ show a systematic tendency to be derived from nouns that designate <state>, <language>, <fundamental>, <system>, and <status>. ¹¹ Similarly, ‘ornative’ verbs tend to stem from nouns indicating <(bio)chemical>, <presentable>, <academic>, <format>, and <rights>.

These generalizations can be formalized into the redundancy rules on vocabulary insertion, as demonstrated in (5) and (6). These rules essentially signify that a noun indicating state or quality and a noun expressing (bio)chemical are inserted under the sister node of $p_{[into]}$ and that of $p_{[with]}$, respectively. Accordingly, the noun *harmony* is correctly inserted into the sister position of $p_{[into]}$, without having to specify that *harmony* is connected to $p_{[into]}$.

(5) $n \rightarrow$ <state>, <language>, <fundamental>, <system>, <status> / $p_{[into]}$ —

(6) $n \rightarrow$ <(bio)chemical>, <presentable>, <academic>, <format>, <rights> / $p_{[with]}$ —

The second subsidiary rule involves the truncation of a word-internal suffix. We have seen in section 2.5 that *-ize* affixation triggers the truncation of an intra-word suffix in the cases of *-ic*, *-ous*, *-ity*, *-ant*, *-ism*, and *-ive* while it may not trigger the truncation in the cases of *-al*, *-(i)an*, *-able*, *-(a)(t)ion*, *-ary*, and *-er*. Moreover, the suffix *-ic* proves to be intact in specific circumstances.

To adjust the morphological structure of *-ize* words, we propose a truncation rule in (7), which is operative in the PF component. ¹² This morpheme-truncation rule entails that *-ic*, *-ous*, *-ity*, *-ant*, *-ism*, and *-ive* are deleted in *-ize* suffixation (cf. *aromatize*) but each of them is not deleted when its base is monosyllabic (cf. *classicize*) or ends in two consonants (cf. *romanticize*). ¹³

(7) *-ic*, *-ous*, *-ity*, *-ant*, *-ism*, *-ive* \rightarrow \emptyset
/ X__ *-ize*

Condition: X=polysyllabic or ending in a single consonant

4 Conclusion

Based on detailed observation of the derived verbs discerned in a large-scale corpus, we have revealed the essential properties of verb derivation. Semantically, derived verbs are divided into five main groups and each submeaning is correlated with a base’s meaning. Formal restrictions are placed on the internal structures and vocabulary strata of the bases. As regards productivity, *-ize* affixation is creative in its construction of numerous innovated verbs. The above properties of derived verbs are theoretically accounted for; basic features common to all five submeanings follow naturally from a core part of their underlying structures. The productivity of *-ize* derivation also arises from its underlying syntactic configuration. Finally, formal restrictions on the bases and the base-derivative meaning correlation originate in the insertion conditions of vocabulary items.

A rigorous analysis of the formal restrictions and the semantic correlation awaits further investigation. Hopefully, we have shown that the study of word formation mechanism can be widely promoted by “corpus-based investigation.”

Acknowledgments

I would like to express my gratitude to three anonymous reviewers for their valuable comments and suggestions on an earlier draft of this paper. This work is partly supported by a Grant-in-Aid for Scientific Research (C) (No. 22K00562) from the Japan Society for the Promotion of Science.

References

- Mark Aronoff. 1976. *Word Formation in Generative Grammar*. MIT Press, Cambridge, MA.
- Harald R. Baayen and Antoinette Renouf. 1996. Chronicling *the Times*: productive lexical innovations in an English newspaper. *Language*, 72:69-96.
- Eve V. Clark and Herbert H. Clark. 1979. When nouns surface as verbs. *Language*, 55:767-811.
- David Embick. 2010. *Localism versus Globalism in Morphology and Phonology*. MIT Press, Cambridge, MA.

¹¹Angle brackets are used here for referring to semantic categories; <language> is intended to mean ‘something related to language.’

¹²See Aronoff (1976: 88-98) for arguments for truncation rules.

¹³Exceptional cases which do not seem to follow the rule are specified on an item-by-item basis. For example, that *-ic* in *aestheticize* is untruncated is specified in the lexical entry of *aesthetic*.

- Joseph E. Emonds. 2000. *Lexicon and Grammar: The English Syntacticon*. Mouton de Gruyter, Berlin.
- Ken Hale and Samuel J. Keyser. 1993. On Argument Structure and the Lexical Expression of Syntactic Relations. In K. Hale and S. J. Keyser, editors. *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, pages 53-109. MIT Press, Cambridge, MA.
- Morris Halle and Alec Marantz. 1994. Some key features of Distributed Morphology. *MIT Working Papers in Linguistics*, 21:275-288.
- Heidi Harley. 2009. Compounding in Distributed Morphology. In R. Lieber and P. Štekauer, editors. *The Oxford Handbook of Compounding*, pages 129-144. Oxford University Press, Oxford.
- Heidi Harley and Ralf Noyer. 2000. Formal versus Encyclopedic Properties of Vocabulary: Evidence from Nominalisations. In B. Peeters, editor. *The Lexicon-Encyclopedia Interface*, pages 349-374. Elsevier, Amsterdam.
- Jennifer Hay. 2003. *Causes and Consequences of Word Structure*. Routledge, New York.
- Otto Jespersen. 1949. *A Modern English Grammar on Historical Principles*, volume 6. George Allen and Unwin, London.
- Beth Levin and Malka Rappaport Hovav. 1995. *Unaccusativity: At the Syntax-Lexical Semantics Interface*. MIT Press, Cambridge, MA.
- Rochelle Lieber. 2005. English Word-Formation Processes. In P. Štekauer and R. Lieber, editors. *Handbook of Word-Formation*, pages 375-427. Springer, Dordrecht.
- Alec Marantz. 1996. 'Cat' as a phrasal idiom: consequences of late insertion in Distributed Morphology. ms., MIT.
- Hans Marchand. 1969. *The Categories and Types of Present-Day English Word-Formation: A Synchronic-Diachronic Approach*, 2nd ed. C. H. Beck, München.
- Ora Matushansky and Alec Marantz. 2013. *Distributed Morphology Today: Morphemes for Morris Halle*. MIT Press, Cambridge, MA.
- Ingo Plag. 1999. *Morphological Productivity: Structural Constraints in English Derivation*. Mouton de Gruyter, Berlin.
- Randolph Quirk, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. Longman, London.
- Sally Rice and Gary Prideaux. 1991. Event-packing: the case of object incorporation in English. *BLS*, 17:283-298.
- Elisabeth O. Selkirk. 1982. *The Syntax of Words*. MIT Press, Cambridge, MA.