



HAL
open science

The influence of nominal etymology on the morphology of English denominal verbs

Aurélie Héois

► **To cite this version:**

Aurélie Héois. The influence of nominal etymology on the morphology of English denominal verbs. *Anglo-phonica*, French Journal of English Studies, 2025, 40, <10.4000/15uyr>. <hal-05578647>

HAL Id: hal-05578647

<https://hal.science/hal-05578647v1>

Submitted on 3 Apr 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

La version éditeur publiée de cet article est disponible en accès libre à l'adresse suivante : <https://journals.openedition.org/anglophonia/7603>

The influence of nominal etymology on the morphology of English denominal verbs

Aurélie Héois, Université Jean Moulin Lyon 3, Centre d'Études Linguistiques – Corpus, Discours et Sociétés, aurelie.heois1@univ-lyon3.fr

RÉSUMÉ

À travers l'analyse d'un échantillon composé de 600 verbes dénominaux anglais, cet article étudie la question de l'influence de l'étymologie et de la morphologie de la base nominale sur la morphologie du verbe dérivé. Des études précédentes ont montré que certaines variables liées à la base nominale ont, dans une certaine mesure, une influence sur la morphologie du verbe. Ce serait notamment le cas de bases empruntées aux langues romanes ou construites par composition. Cependant, ces mêmes études indiquent aussi que certaines variables ne seraient plus pertinentes pour expliquer la morphologie des verbes dénominaux en anglais contemporain. Comme les données présentées s'étalent sur deux siècles (1800-2022) et incluent la période contemporaine, la pertinence de ces variables est examinée pour toute la période. Des tests statistiques sont réalisés sur les données pour déterminer si l'influence des différentes variables est significative. L'étude confirme certaines hypothèses préalables et propose de nouvelles contraintes pour les procédés de formation lexicale moins étudiés dans les travaux sur les verbes dénominaux, à savoir la rétroformation, la troncation, la parasynthèse et la dérivation paradigmatique. L'article conclut en présentant des données chronologiques qui suggèrent un déclin dans la dérivation des verbes dénominaux sur des bases empruntées à d'autres langues et ce depuis le début du 20^{ème} siècle. Cette observation pourrait expliquer en partie le constat d'un affaiblissement de l'influence de l'étymologie romane sur la morphologie verbale.

ABSTRACT

Through the analysis of a sample of 600 denominal verbs in English, this article investigates the debate around the influence of nominal etymology and morphology on the morphology of the verbal output. Previous studies have suggested that base nouns variables, such as borrowing from Romance languages or compounded constructions, have, to a certain extent, an influence on the morphology of the derived verb. However, these studies also indicate that some variables may not be relevant in present-day English any longer. As the dataset under scrutiny covers two centuries (1800-2022) and incorporates the most recent period, the relevance of these variables is questioned for the whole period. Statistical tests of significance are carried out in order to assess this supposed influence. The study confirms some of the previous hypotheses and proposes new constraints for lesser researched word-formation processes, such as back-formation, clipping, parasynthesis and paradigmatic derivation. It concludes on chronological data which suggest that the construction of English denominal verbs on borrowed bases has declined since the beginning of the 20th century which may help understand why the influence of Romance etymology is considered more and more anecdotal.

Mots-clés : verbe dénominal, morphologie, étymologie, origine, influence.

Key words: denominal verb, morphology, etymology, origin, influence.

1. Introduction

There seems to be a general consensus in the literature on denominal verbs that the etymology of the base noun, its origin, has an impact on the morphology of the derived verb (Plag, 1999; Dixon, 2008; Debouzie, 2024). The extent of this influence, however, is not always clear. Dixon (2008: 32) claims that “the motivating factors [which explain the morphosemantics of the verb] are a combination of phonological form, semantics and etymology”. In this perspective, *etymology* reflects

the geographical origin of a word, whether it is borrowed from another language or not, and more generally to which language family it originally belongs. The definition of *etymology* adopted in this study is broader as it includes both the geographical origin and the morphological makeup of a word. Regarding the geographical origin of the base noun, Plag (1999: 58) is more nuanced and distinguishes between affixes with restrictions, such as *-en* which forms verbs from native inputs only, or Latinate suffixes, such as *-ize*, which “primarily attach to Latinate stems” and other types of affixes which are not influenced by the origin of the base noun. Similarly, Haspelmath and Sims (2010: 122) suggest that some English suffixes are restricted to specific types of bases but that verbal suffixes such as *-ize* and *-ify* “seem to have lost their restriction to Latinate bases almost entirely”. Nonetheless, Debouzie (2024: 198) finds a positive correlation between Latinate base nouns and affixation for contemporary English denominal verbs (first attested after 1950).

There are two main types of geographical origins: the base noun is either endogenous to English and has developed within the language through various morphological processes, or it was borrowed from another language and is thus exogenous. This dichotomy is sometimes referred to as native vs. non-native (Bauer *et al.*, 2013: 35). In English, borrowed bases are often Latinate, mostly originating from French or Latin. The term *native* is frequently replaced by *Germanic* although the latter is broader as it encompasses both native English lexemes and borrowings from other Germanic languages. In the present study, endogenous, or native, bases are restricted to the English language. They include nouns which are reflexes in present-day English of Old or Middle English words, such as *tail*, as well as nouns which have undergone word-formation processes within the English language, as in *childminder*.

While most studies agree that etymology influences the morphology of the derived verb, the extent of this influence varies. Moreover, in most cases, only what are considered the most productive processes are taken into account. As a result, many questions remain unanswered. Thus, assessing the weight of etymological parameters on the morphology of denominal verbs is still relevant. This leads to the research question: how significant is the influence of the etymology of the base noun on the morphology of the derived verb? The present study focuses on a dataset of denominal verbs first attested between 1800 and 2022 with the aim of systematically analyzing the relationship between the etymology of the base noun and the morphology of the derived verb, including all morphological processes available for creating denominal verbs.

In the next section, I present an overview of the various hypotheses made on the influence of the origin of the base noun on the morphology of the verb. In section 3, I first explain how the data was collected and selected for this study before presenting the derived verbs according to two criteria, viz. origin of the base noun and morphology of the verb. The fourth section focuses on the bidimensional analyses of the data along with statistical tests of significance. The last section is devoted to the discussion of these results and what they bring to the debate on the influence of the nominal base on the morphology of the verb.

2. Influence of the etymology of the base noun on the morphology of denominal verbs: an overview

Many studies on denominal verbs restrict their analyses to some specific word-formation processes. For instance, the generative tradition tends to restrict the definition of denominal verbs to conversion only (Kiparsky, 1997; Bleotu, 2019). For other reasons, Clark and Clark (1979) only explore cases of conversion as well. More comprehensive studies on denominal verbs usually include the two main word-formation processes which are considered productive in denominal verbal derivation (Plag, 1999: 117), namely affixation and conversion (Plag, 1999; Dixon, 2008; Debouzie 2024). A denominal verb is understood as a verb originating from a noun and sharing both morphological and semantic elements with this noun. Hence, the present study proposes to explore the full set of word-formation processes that can come into play in the construction of denominal

verbs: affixation, conversion, back-formation, clipping as well as more complex processes (see section 2.3). The following section gives an overview of the debate on the influence of nominal origin on the derived verb along with the results which might be expected in this study.

2.1. Influence of nominal etymology in conversion

Conversion is a word-formation process in which an input and an output share the same form, are semantically related, but belong to different grammatical categories (Debouzie, 2024: 18; among others). According to Plag (1999: 117), conversion is the most productive word-formation process which can create a denominal verb in modern English. While Bauer *et al.* (2013: 277) “generally support this conclusion”, they add some nuance by pointing out that methodological issues make it difficult to compare the productivity of the different word-formation processes. Regarding the types of bases that can be selected in the process of conversion, Bauer (1983: 226) states that it is available to all types of lexemes without almost any restriction:

There do not appear to be morphological restrictions on the forms that can undergo conversion, so that compounds, derivatives, acronyms, blends, clipped forms and simplex words are all acceptable inputs to the conversion process. Similarly, all form classes seem to be able to undergo conversion [...].

As such, there should not be any strong correlation between the etymology of the noun and the process of conversion. Bauer *et al.* (2013: 278) list various examples of nouns which have been converted into verbs: simplex bases, either endogenous to English, such as *eel*, or borrowed from another language, such as *archive* from French; suffixed nouns such as *junction* or *sleeper*; compounds as in *eyeball* or *postcard*; as well as proper nouns (*MacGyver*) and acronyms (*RIF*).

Some studies mention, however, elements which can either restrict or favor conversion. It has been noted that if the base noun originates from the suffixation of a verb, then the output is blocked by the existence of the already existing verb (Bauer, 1983: 226; Marchand, 1969) unless the meaning of the new verb is different from the meaning of the previous verb (Bauer, 1983: 227). For instance, on the one hand, Marchand (1969) signals that the conversion of *arrival*_N is blocked by the verbal base *arrive*_V; on the other hand, in case of semantic differentiation, the deverbal nominal suffix does not block conversion, as is the case of *signal*_V (< *signal*_N < *sign*_V < *sign*_N) (Bauer, 1983: 227). If there is such a restriction on conversion, the data will show that suffixed base nouns are usually favored by processes other than conversion.

Plag (1999: 232) shows that compounded base nouns, such as *highlight*, tend to favor conversion as their complex morphology is expected to block any process of suffixation. Bauer *et al.* (2013: 510) give a list of all the affixes which do not combine with compound bases. This list includes all the verbalizing suffixes *-ize-*, *-ify*, *-ate*, *-en*, along with verbalizing prefixes *be-*, *dis-* and *en-*. Consequently, a correlation is likely to be found between compounds and processes other than affixation, especially with conversion as it is the most productive word-formation process in English.

Additionally, Dixon (2008: 47) claims that “[t]ypically, a Germanic noun will form a verb by zero derivation, while a Romance loan of similar meaning will take *-ize*”. However, he also states earlier in his study that conversion contains the same proportion of verbs based on Latinate nouns as it does of verbs based on Germanic nouns (Dixon, 2008: 38). Therefore, the expected observation will be of a relative overrepresentation¹ of Germanic bases in converted verbs while a similar result will be expected for Latinate bases and suffixation.

Finally, while most studies focus on Latinate and Germanic bases, which account for the majority of the English lexicon, the question of base nouns borrowed from other language families remain. Debouzie (2024: 198) suggests that a significant correlation exists between base nouns whose origin is neither Latinate nor Germanic and conversion.

2.2. Influence of nominal etymology in affixation

There are two types of affixations involved in the derivation of denominal verbs: suffixation and prefixation. Prefixation is a process by which a prefix attaches before the base, while suffixation involves a suffix which attaches after the base (Bauer, 2014: 124).

According to Plag (1999: 117), suffixation in *-ize* is the second most productive process to form denominal verbs in English. Other types of suffixations include *-ify*, *-ate* and *-en*. Plag (1999: 70) observes that only the native suffix *-en*, as in *crispen* (Plag, 1999: 219), shows an etymological restriction to base nouns of native origin. This observation is consistent with Haspelmath and Sims's view that the Latinate constraint has been lost for *-ize* and *-ify* in modern English (2010: 122). However, Dixon (2008: 38), who does not solely focus on modern English, holds that "the great majority of [the nouns] verbalized through *-(i)fy* or *-ize* (or suffix *-ate* or prefix *en-*) are of Romance origin". Bauer *et al.* (2013: 269) add more nuance to this observation stating that non-native bases only "predominate" in these two paradigms. As a consequence, it is unclear whether the Latinate constraint is relevant for data ranging over two centuries (1800-2022), which is why it is investigated in this study.

Plag (1999: 82) notes that the *-ate* suffix is "restricted to chemical substances, whose names [...] usually do not involve suffixes". As a result, the expected observations for *-ate* denominal verbs should suggest that suffixed base nouns are usually not involved in this type of suffixation. By contrast, "[t]he denominal suffix *-ize* seems to attach quite often, and naturally, to suffixed nouns of various types" (Plag, 1999: 83), so that an overrepresentation of suffixed bases is anticipated in this case.

Most studies seem to agree that prefixation is not very productive in English denominal verbs. Dixon (2008: 32) notes only two prefixes: *en-* (and its allomorph *em-*) of Romance origin, as in *envision_V*; and *be-* of Germanic origin, as in *behead_V*. He supposes a Latinate constraint on the first prefix (Dixon, 2008: 38) but Bauer *et al.* (2013: 268) note that this prefix can be found on both native and non-native bases, as in *encash* where the noun *cash* is either borrowed from French or Italian (*OED*, 1888), or in *enhat* where *hat* is of Germanic origin (*OED*, 2017). They make a similar observation for *be-*. Bauer *et al.* (2013: 267, 343, 358) as well as Debouzie (2024: 59-60) add four verbalizing prefixes to the list: the native prefixes *out-* and *un-*, as in *unearth* and *out-birth*; and the non-native prefixes *de-*, as in *dethrone*, and *dis-*, as in *disrobe*. Debouzie (2024) notes that the native prefix *un-* usually attaches to native bases.

Contrary to most affixes, Bauer *et al.* (2013: 511-513) note that prefixes *de-* and *out-* may occasionally combine with compounded bases while prefix *un-* frequently does so, although it is unclear from their examples if that is the case for verbalizing prefixes.

2.3. Influence of nominal etymology in other types of word-formation processes

As most studies focus on major word-formation processes, little has been said about the influence of the etymology of the base noun on the morphology of the verb for other processes involved in denominal verbal derivation, namely back-formation, clipping and complex processes.

Back-formation is the "deletion of a suffix or prefix in analogy to pairs of base and derivative that feature the affix in question" (Bauer *et al.*, 2013: 20). The pair *burgle* < *burglar* is a famous example of noun-to-verb back-formation.

Similarly, clipping is a type of subtractive word-formation process as it involves the "deletion of some of its phonological material" as in the noun-to-noun example *lab* (from *laboratory*) (Bauer *et al.*, 2013: 25). The main difference between clipping and back-formation lies in the nature of the deleted material. Back-formation involves the deletion of material which can be morphologically

analyzed as an affix, as in example (1), but what is deleted in clipping cannot, as illustrated in (2) (examples from VdenomEN, see section 3.1).

(1) buttle: “to pour out (a drink)” (*OED*, 2018) < butler

(2) hool: “to drive fast (and potentially dangerously)” (*GDS*, 2022) < hooligan

The label *complex processes* is used here to categorize cases which do not fit in the previously defined categories, as illustrated in examples (3) to (5) (examples from VdenomEN, see section 3.1). In most cases, they can be analyzed as the simultaneous combination of two word-formation processes. Parasynthesis is the simultaneous combination of a prefix with a suffix (Bauer *et al.*, 2013: 500), as illustrated in (3). Paradigmatic derivation, also called *crossformation* (Haspelmath and Sims, 2010: 49) or *paradigmatic morphological process* (Bauer *et al.*, 2013: 20) is a process which is based on a morphological paradigm which can be found regularly in the lexicon, as illustrated in (4) for the *-ism/-ize* paradigm. This process could also be analyzed as the combination of back-formation with suffixation, although the use of the term “paradigmatic” makes it possible to avoid a linear analysis. In cases where no regularity can be found, I resort to this latter analysis for the purpose of notation. Hence, example (5) is coded as the combination of the deletion of *-ition* (back-formation) with a suffixation in *-ize*.

(3) disinsectize: “to remove insects from, esp. from an aircraft” (*OED*, 1972)ⁱⁱ < insect

(4) metasomatize: “to alter (a rock, mineral, etc.) by metasomatic processes” (*OED*, 1942) < metasomatism

(5) premunize: “To introduce pathogens into so as to produce premunition” (*OED*, 2007) < premunition

The lack of data on the influence of the origin of the base noun for these processes either suggests that the question is not relevant or that no systematic study has yet been conducted on this issue. The only hypothesis that can be drawn concerns parasynthesis, paradigmatic derivation and other complex processes involving affixation. Indeed, it is possible that the same restrictions or attractions found for the affixes also apply in these cases.

2.4. Summary of the expected results

Table 1 summarizes the different results that can be expected according to the previous studies presented above. Each result will be assessed in light of the analysis of data ranging from 1800 to 2022. Some findings from previous studies may not be confirmed by our analysis as the periods under consideration vary. The main input of the present study lies in its inclusion of all derivation processes that can be found in the formation of English denominal verbs.

Word-formation processes	Expected results
Conversion	<ul style="list-style-type: none"> ▪ almost no restriction ▪ overrepresentation of Germanic base nouns ▪ overrepresentation of nouns borrowed from languages other than Germanic and Latinate languages ▪ overrepresentation of compound bases ▪ underrepresentation of suffixed base nouns
Suffixation in <i>-ate</i>	<ul style="list-style-type: none"> ▪ overrepresentation of Latinate bases ▪ underrepresentation of suffixed bases ▪ no compound bases
Suffixation in <i>-en</i>	<ul style="list-style-type: none"> ▪ only Germanic bases

	<ul style="list-style-type: none"> ▪ no compound bases
Suffixation in <i>-fy</i>	<ul style="list-style-type: none"> ▪ overrepresentation of Latinate bases ▪ no compound bases
Suffixation in <i>-ize</i>	<ul style="list-style-type: none"> ▪ overrepresentation of Latinate bases ▪ overrepresentation of suffixed bases ▪ no compound bases
Prefixation in <i>be-</i>	<ul style="list-style-type: none"> ▪ no preference on the origin of the base ▪ no compound bases ▪ overrepresentation of suffixed bases
Prefixation in <i>de-</i>	<ul style="list-style-type: none"> ▪ no preference on the origin of the base ▪ occasional compound bases ▪ overrepresentation of suffixed bases
Prefixation in <i>dis-</i>	<ul style="list-style-type: none"> ▪ no preference on the origin of the base ▪ no compound bases ▪ overrepresentation of suffixed bases
Prefixation in <i>en-</i>	<ul style="list-style-type: none"> ▪ no preference on the origin of the base ▪ no compound bases ▪ overrepresentation of suffixed bases
Prefixation in <i>out-</i>	<ul style="list-style-type: none"> ▪ no preference on the origin of the base ▪ occasional compound bases ▪ overrepresentation of suffixed bases
Prefixation in <i>un-</i>	<ul style="list-style-type: none"> ▪ mostly native bases ▪ frequent compound bases ▪ overrepresentation of suffixed bases
Back-formation	<ul style="list-style-type: none"> ▪ overrepresentation of suffixed bases ▪ overrepresentation of compound bases
Clipping	<ul style="list-style-type: none"> ▪ overrepresentation of suffixed bases ▪ overrepresentation of compound bases
Complex processes	<ul style="list-style-type: none"> ▪ same restrictions and attractions as affixes

Table 1. Expected influence of the etymology of the base noun on the morphology of the verb

3. Methodology

3.1. Data collection and selection: VdenomENⁱⁱⁱ

The VdenomEN database has been created in order to study English denominal verbs first attested between 1800 and 2022^{iv}. Thanks to the advanced search tool available for both the *Oxford English Dictionary online (OED)* and *Green's Dictionary of Slang (GDS)*, I obtained a list of all recorded verbs first attested between 1800 and 2022. I then manually selected all denominal verbs from these two lists. In total, I gathered 6 933 English verbs. In order to carry out the analysis on a workable sample, I randomly selected 600 of those verbs. I then coded three variables for the origin of the base noun and two variables to describe the morphology of the verb.

As explained in the introduction, the notion of *etymology* encompasses both the geographical origin of the base noun and, for endogenous bases, their morphological makeup. As a result, three variables were coded to account for the origin of the base noun. The first origin variable, called “Origin1_N”, is the most coarse-grained and takes 4 different and mutually exclusive values: “B” for borrowings, “EN” for native nouns (“EN” stands for English), “PN” for proper names, and “?” when the origin of the base noun is unknown. As the aim of the study is to explore whether there is a significant link between the origin of the base noun and the morphology of the verb, I decided to exclude two values in this first variable, “PN” and “?”. The latter does not need further details, but the former may need some explanations. Some denominal verbs, such as *Napoleonize (OED, 2003)*^v,

are built on proper names. However, deciding on the etymology of a proper name can be problematic and, more specifically, deciding on which part of the etymology is relevant for the study of denominal verbs is an issue. For instance, some proper names have a morphologically transparent etymology that is endogenous, that is to say built from English words. *Bluegrass State*^{vi} (GDS, 2022) is such a case and its base noun could be coded “EN”. However, it is harder to decide for other proper names. For example, the verb *Napoleonize*^{vii} (OED, 2003) is built on the proper name *Napoleon*, the name of the French emperor pronounced within the English phonological system. Should we consider this proper name a French borrowing, a Corsican or an Italian one? I have chosen to put proper name bases aside for the purpose of this study.

The exclusion of these two values results in a sample of 548 verbs whose base noun is either borrowed from another language (“B”) or endogenous to English (“EN”). As explained in the introduction, nouns are categorized as endogenous in two situations: (1) they are inherited from Old or Middle English; (2) they have been formed within English through morphophonological processes^{viii}. The following section will present the data, first through the origin variables and then through the morphological variables, all of which will be defined through the section.

3.2. Data overview

3.2.1. Origin variables

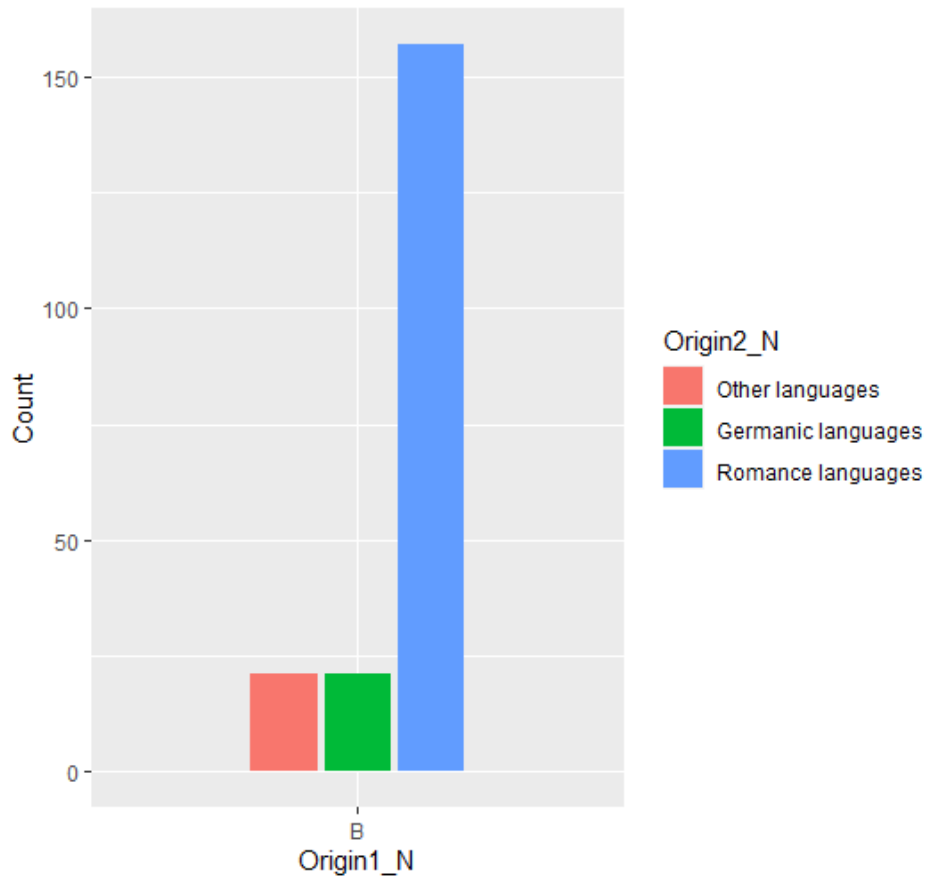
Overall, the data is mostly composed of denominal verbs whose base noun is endogenous to English. 63.7% of denominal verbs in the sample are built on English words while 36.3% come from borrowed base nouns.

For instance, the verb *people-watch*, attested in English in the 1960s, is built on the noun *people-watching*, which itself is formed through the compounding of two nouns (OED, 2005)^{ix}. The verb *ritualize*, on the other hand, whose first meaning is now obsolete, was used in the 19th century, and is built on the noun *ritual* which itself is borrowed from French (OED, 2010)^x.

The second origin variable, called “Origin2_N”, gives more details about the origin of the base noun by specifying if the source language is part of “Germanic languages”, “Romance languages” or “Other languages”. As a result, endogenous base nouns, coded “EN” in the first variable, are part of the “Germanic languages” in this second variable. Other Germanic languages present in the dataset are Afrikaans (*jol*_V < *jol*_N), German (*coalify*_V < *coal*_N), Low German (*unmate*_V < *mate*_N), Dutch (*stillage*_V < *stillage*_N), Middle Dutch (*slip*_V < *slip*_N), Norwegian Nynorsk (*telemark*_V < *telemark*_N), Swedish (*nickelize*_V < *nickel*_N), Yiddish (*gunzel*_V < *gunsel*_N) and unspecified Germanic languages (*gype*_V < *gype*_N). Base nouns borrowed from Romance languages mostly come from French (*ritualize*_V < *ritual*_N) and Latin (*pedestrianize*_V < *pedestrian*_N), but some are also borrowed from Spanish (*disembargo*_V < *embargo*_N), Italian (*microscopize*_V < *microscope*_N) and Portuguese (*palanquin*_V < *palanquin*_N). Finally, the category “Other languages” includes a broad variety of languages: Algonquian (*skunk*_V < *skunk*_N), Australian (*yacker*_V < *yakka*_N) and South American Indian languages (*curarize*_V < *curare*_N), Ancient Greek (*x*_V < *x*_N) and Modern Greek (*calotype*_V < *calotype*_N), Persian (*dakma*_V < *dakhma*_N), Scottish Gaelic (*gulder*_V < *gulder*_N), Gujarati (*tank*_V < *tank*_N), Hindi (*cowage*_V < *cowage*_N), Igbo (*obeah*_V < *obeah*_N), Maori (*moko*_V < *moko*_N), Ojibwa (*totem*_V < *totem*_N), Sanskrit (*pundit*_V < *pundit*_N) and Chinese (*kowtow*_V < *kowtow*_N). The choice to group them in a miscellaneous category is not satisfactory but linked to the lack of examples for these languages.

Figure 1 presents the distribution of borrowed base nouns (excluding endogenous nouns). It shows that 78.9% of the denominal verbs are built on nouns borrowed from Romance languages, mostly Latin and French. When taking endogenous base nouns into account as well, Romance languages still account for 28.6% of the sample, Germanic languages for 67.5% and other languages for 3.8%.

Figure 1. Origin of borrowed base nouns (VdenomEN)



In order to account for the morphological makeup of endogenous nouns, I created a third origin variable, “Origin3_N”. With this variable, borrowed base nouns are coded “B”. Originally, I coded the word-formation processes with some details featuring the parts of speech used to form the noun. For instance, N+N would mean that the noun results from the compounding of two nouns. In order to analyze the data, however, I consolidated the variable and kept only 4 values by grouping morphological makeups according to their type and structure: “COMP” stands for base nouns formed through the compounding of two lexemes, as illustrated in (6) and (7); “SUFF” stands for base nouns built either through suffixation, as illustrated in (8), or through more complex processes involving a suffixation, as in (9); “PREF” is coded for base nouns formed through the addition of a prefix to a base, as in (10); and “SPX” stands for base nouns formed on the basis of a single lexeme, as in (11) and (12). In order to avoid ambiguity between processes linking the base noun to its denominal verb (see section 3.2.2 below) and these processes which describe the morphological makeup of the base noun, I prefixed each value with “MMN” (for “morphological makeup of the noun”).

(6) *password*_V < *password*_N < *pass*_N + *word*_N (MMN_COMP)^{xi}.

(7) *greatcoat*_V < *greatcoat*_N < *great*_A + *coat*_N (MMN_COMP)^{xii}

(8) *reviver*_V < *reviver*_N < *revive*_V + *-er*_{SUFF} (MMN_SUFF)^{xiii}

(9) *metasomatize*_V < *metasomatism*_N < *metasomati*_C + *-ism*_{SUFF} (MMN_SUFF)^{xiv}

(10) *non-intervene*_V < *non-intervention*_N < *non*_{PREF} + *intervention*_N (MMN_PREF)^{xv}.

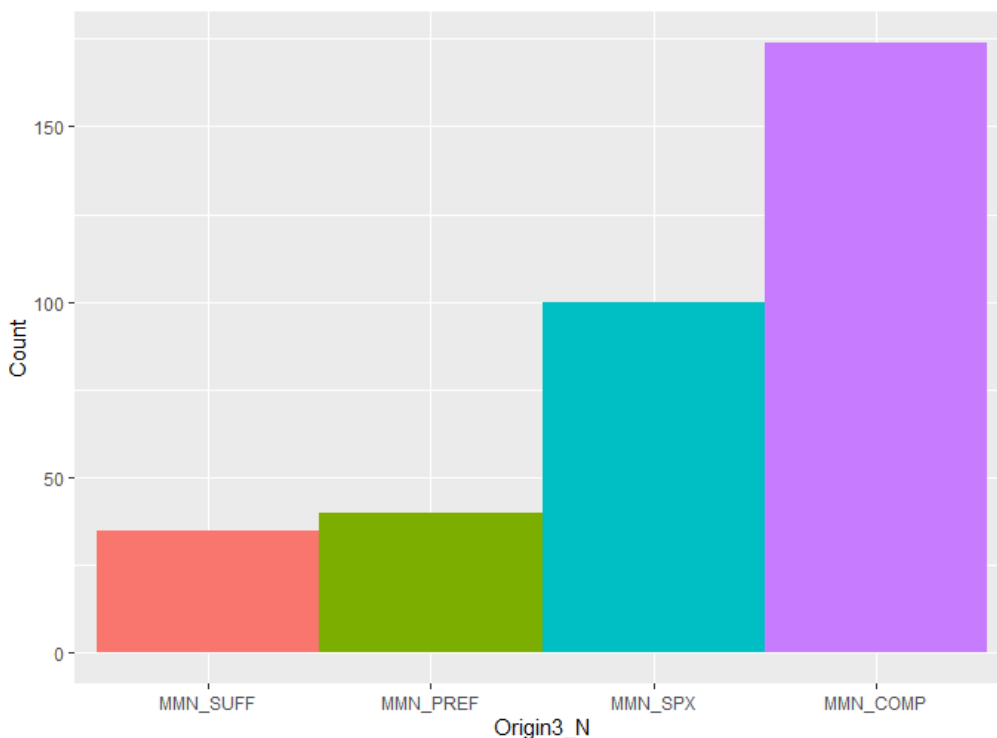
(11) *boob*_V < *boob*_N < *booby*_N (MMN_SPX)^{xvi}.

(12) *weaponize*_V < *weapon*_N < Old English (MMN_SPX)^{xvii}.

In (6), the verb *password* is converted from the noun *password* which is itself formed through the compounding of two nouns *pass* and *word*. *Greatcoat* in (7) is also a noun-to-verb conversion based on a noun formed through the compounding of the adjective *great* with the noun *coat*. The verb *reviver* in (8) is also converted but its base noun is formed through *-er* suffixation on the verb *revive*. *Metasomatize*, in (9), involves paradigmatic derivation on three levels: a verb in *-ize*, based on a noun in *-ism*, itself based on an adjective in *-ic*; because the base noun involves a suffix, it is categorized as MMN_SUFF. In (10), the verb *non-intervene* is based on the noun *non-intervention* which is itself formed through a process of prefixation on a nominal base. The verb *boob* in (11) is an example of MMN_SPX: the verb is a conversion from a noun which itself is the result of a back-formation from the noun *booby*. Also categorized as MMN_SPX, *weaponize* in (12) is a verb formed through the suffixation of the noun *weapon* which, according to the *OED* (1926), is a “word inherited from Germanic”. As the etymology section shows the evolution of the word from pre-Germanic to Old English, the present-day form of the word is considered to be a reflex of an Old English lexeme.

Figure 2 shows the distribution of “Origin3_N” in the sample (the graph only includes endogenous base nouns, however the percentages given are for the whole sample, that is to say both borrowed and endogenous base nouns). Most endogenous base nouns in the sample have a compounded structure (“MMN_COMP”, 31.8%). 18.2% of base nouns in the sample are formed through a process which includes a single lexeme and no affix (“MMN_SPX”) while 7.3% are prefixed (“MMN_PREF”) and 6.4% are suffixed (“MMN_SUFF”).

Figure 2. Morphological makeup of endogenous base nouns (VdenomEN)

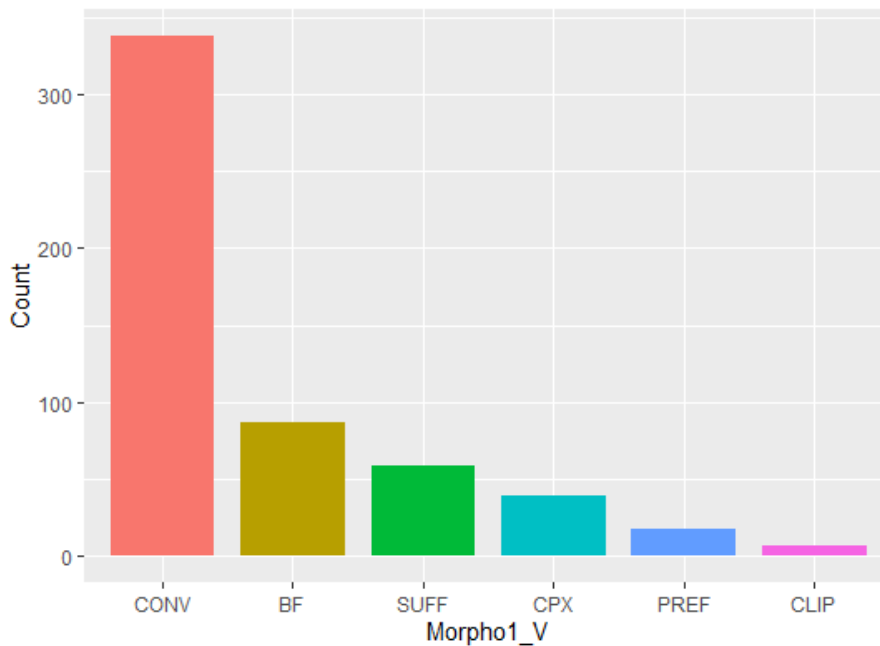


3.2.2. Morphological variables

Two morphological variables are coded for the verbal output. They vary in terms of descriptive accuracy. The first variable, “Morpho1_V”, is coarse-grained and includes 6 values which describe the 6 word-formation processes available to form denominal verbs in English (see section 2): in the sample, 61.7% are formed through conversion (“CONV”), 15.9% through back-formation (“BF”), 10.8% through suffixation (“SUFF”), 7.1% through complex processes (“CPX”, encompassing parasynthesis, paradigmatic derivation and back-formation + affixation), 3.3% through prefixation

(“PREF”) and 1.3% through clipping (“CLIP”). Figure 3 presents the distribution of the data according to this first morphological variable.

Figure 3. Verbal morphology (VdenomEN)



The second morphological variable, “Morpho2_V”, is finer-grained and comprises 14 values. Some processes, such as conversion and clipping, remain the same between the two variables and examples are given in (13) and (26), while other processes are detailed in the second variable. Back-formation is subdivided into four values depending on the deleted element: “BF_ion”, “BF_ing”, “BF_er” and “BF_OTH”. “BF_ion” groups together verbs formed through the deletion of *-ion*, *-ation*, *-cation* and *-tion*, as illustrated in (14); “BF_ing” and “BF_er” categorize verbs built from nouns through the deletion of *-ing* and *-er* respectively, examples are given in (15) and (16); “BF_OTH” classifies other cases of back-formation which correspond to 1, 2 or 3 instances in the sample, some examples are given in (17). Five types of suffixation can be found in the sample but only the suffix *-ize* is frequent in the data (40 occurrences). As the other suffixes only account for less than 10 instances, the hypotheses formulated in section 2 cannot be checked individually for these affixes. As a result, the variable takes two values for suffixation: “SUFF_ize” and “SUFF_OTH”. Both are illustrated with examples in (18) and (19). Complex processes, “CPX”, encompass three types of processes: parasynthesis, paradigmatic derivation and what is analyzed as a combination of back-formation and affixation (without any theoretical assumption). The common feature of all these processes is the presence of affixation, suffixation in most cases. As explained in section 2, for complex processes, the only hypothesis that can be drawn from previous studies lies in the comparison with other processes, namely affixation. As a result, these verbs are categorized along three values according to the affix they include: “CPX_ize” groups together all verbs whose formation involves *-ize*, examples are given in (20); “CPX_SUFF” categorizes verbs displaying another type of suffix, as in (21); “CPX_PREF” groups together verbs involving prefixation without suffixation, examples are given in (22). There are only 18 examples of prefixation in the sample involving 7 different prefixes: *be-*, *de-*, *dis-*, *em-/en-*, *in-*, *re-* and *un-*. According to previous studies, only *un-* shows selection restrictions (native bases only). Moreover, there appears to be a difference between prefixes when it comes to compounded bases: some never appear with compounded bases (*be-*, *dis-*, *en-/em-*), while others may attach to compounded bases either occasionally (*out-*, *de-*) or frequently (*un-*). Two of the prefixes found in the sample were not listed as verbalizing prefixes earlier: *re-* and *in-*. The prefix *re-* is usually added to a verb base and does not appear to show any selection restriction on the bases it can attach to (Bauer *et al.*, 2013: 419). The prefix *in-*, in its locative meaning, can attach to nouns, verbs and adjectives and Bauer *et al.* (2013: 342) suggest that it does

not attach to compounded bases. As a result, I chose to subdivide prefixes into three categories: “PREF_un” includes all verbs formed through the prefixation of *un-*, it is expected to attach to native bases only and to be compatible with compounded bases; “PREF_comp” groups together verbs formed through the prefixation of *de-* and *re-*, they are expected to be compatible with compounded bases; “PREF_nocomp” encompasses prefixations in *be-*, *dis-*, *en-/em-* and *in-* and is expected to be incompatible with compounded bases. Examples (23) to (25) illustrate the three categories.

- (13) CONV: *revolver_v* < *revolver_N*, *code name_v* < *code name_N*, *spasm_v* < *spasm_N*, *grasshopper_v* < *grasshopper_N*
- (14) BF_ion: *cryopreserve_v* < *cryopreservation_N*, *remediate_v* < *remediation_N*, *sorb_v* < *sorption_N*
- (15) BF_ing: *market test_v* < *market testing_N*, *headbang_v* < *headbanging_N*, *type-cast_v* < *type-casting_N*
- (16) BF_er: *buttle_v* < *butler_N*, *steamroll_v* < *steamroller_N*, *lech_v* < *lecher_N*
- (17) BF_OTH: *amanuense_v* < *amanuensis_N*, *chiromance_v* < *chiromancy_N*, *abey_v* < *abeyance_N*
- (18) SUFF_ize: *morselize_v* < *morsel_N*, *weaponize_v* < *weapon_N*, *rivalize_v* < *rival_N*, *structurize_v* < *structure_N*
- (19) SUFF_OTH: *piecen_v* < *piece_N*, *shinny_v* < *shin_N*, *coalify_v* < *coal_N*, *acetylate_v* < *acetyl_N*
- (20) CPX_ize: *intelligize_v* < *intelligence_N*, *disinsectize_v* < *insect_N*, *eulogize_v* < *eulogy_N*
- (21) CPX_SUFF: *denaturate_v* < *nature_N*, *deconstructify_v* < *destruction_N*, *maximate_v* < *maximum_N*
- (22) CPX_PREF: *decrypt_v* < *cryptogram_N*, *dispageant_v* < *pageantry_N*
- (23) PREF_un: *unurn_v* < *urn_N*, *unmate_v* < *mate_N*
- (24) PREF_comp: *delouse_v* < *louse_N*, *re-rail_v* < *rail_N*
- (25) PREF_nocomp: *betail_v* < *tail_N*, *discaste_v* < *caste_N*, *enhat_v* < *hat_N*, *embus_v* < *bus_N*, *intrine_v* < *trine_N*
- (26) CLIP: *goo-goov* < *goo-goo eyes_N*, *tish_v* < *tissu_N*

Figure 4 presents the distribution of the second morphological variable for back-formation, suffixation, complex processes and prefixation. It is organized by word-formation type in decreasing frequency.

Figure 4. Zoom on verbal morphology
back-formation, suffixation, prefixation and complex processes (VdenomEN)



The most frequent processes in the sample are: *ion*-back-formation (7.8%), *ize*-suffixation (7.3%) and complex processes involving *ize*-suffixation (4.9%). Other types of suffixation account for 3.5% of the sample, *ing*-back-formation for 3.3%, *er*-back-formation for 2.4%, other types of back-formation for 2.4%, complex processes involving suffixes other than *-ize* for 1.6%, prefixation not compatible with compounded bases for 1.5%, prefixation compatible with compounded bases for 1.1%, prefixation in *un-* for 0.7% and complex processes involving prefixation but no suffixation for 0.5%.

4. Results: bidimensional analyses

In order to determine whether there is a correlation between the origin of the base noun and the morphology of the verb, a bidimensional analysis has been carried out. Each origin variable has been paired with a morphological variable. As there are three origin variables and two morphological variables, this analysis amounts to six pairs. For each pair of variables, I proceeded as follows: I first created a contingency table; and then I carried out the chi-squared test and the Pearson's residuals test using RStudio (in the tables below, each cell presents first the line proportion and then Pearson's residual). As the results of statistical analyses are strongly dependent on selected variables and the way they are built, I chose to pair all variables in order to obtain a clearer picture of the influence of nominal etymology on the morphology of the derived verb.

4.1. Pair 1: Origin1_N and Morpho1_V

Origin1_N	EN	B	Total
Morpho1_V			
CONV	68.3 1.07	31.7 -1.42	100
BF	72.4 1.02	27.6 -1.35	100
SUFF	39.0	61.0	100

	-2.38	3.15	
CPX	38.5 -1.97	61.5 2.61	100
PREF	61.1 -0.14	38.9 0.18	100
CLIP	85.7 0.73	14.3 -0.97	100
Total	63.3	36.7	100

Table 2. Contingency table and Pearson's residuals for Morpho1_V and Origin1_N (p<0.05)

For the pair of variables Morpho1_V and Origin1_N, the p-value shows a significant correlation between the two, suggesting that there might be a significant influence of the origin of the base noun on the morphology of the verb. However, the picture is not as clear-cut as the p-value suggests and the residuals show a more nuanced influence. Only three cells in Table 2 show significant results ($> |2|$, in bold). According to these data, suffixation favors borrowed bases while it disfavors endogenous bases. The third significant correlation concerns complex processes which seem to favor borrowed bases as well. As complex processes usually involve suffixation, this correspondence between complex processes and borrowings may be linked to the correlation between suffixation in general and borrowed bases.

4.2. Pair 2: Origin2_N and Morpho1_V

Origin2_N Morpho1_V	Germanic languages	Romance languages	Other languages	Total
CONV	72.2 2.96	22.8 -3.85	5.0 1.85	100
BF	72.4 1.06	26.4 -0.50	1.1 -1.42	100
SUFF	49.2 -3.19	50.8 3.99	0.0 -1.62	100
CPX	41.0 -3.67	51.3 3.24	7.7 1.30	100
PREF	66.7 -0.08	33.3 0.45	0.0 -0.86	100
CLIP	85.7 1.03	14.3 -0.85	0.0 -0.53	100
Total	67.5	28.6	3.8	100

Table 3. Contingency table and Pearson's residuals for Morpho1_V and Origin2_N (p<0.05)

For the pair of variables Morpho1_V and Origin2_N, the p-value shows a significant correlation but, as with the previous variables, this apparent general correlation hides a more complex picture. The Pearson's residuals show that there is no significant relationship between other languages and morphology. This can probably be explained by the lack of data for this category and further research is needed to reach any conclusion. The data however shows significant tendencies for Romance and Germanic languages. There is a positive correlation between Romance languages and suffixation as well as complex processes, suggesting that, as base nouns, Romance lexemes are overrepresented in these two types of processes. Conversely, Romance languages are underrepresented in conversion. The patterns are reversed for Germanic bases as the residuals show a positive correlation with conversion and a negative correlation with both suffixation and complex processes.

4.3. Pair 3: Origin3_N and Morpho1_V

Origin3_N Morpho1_V	MMN_ SUFF	MMN_ PREF	MMN_ SPX	MMN_ COMP	B	Total
CONV	4.1 -1.63	4.1 -2.15	21.3 1.31	38.8 2.29	31.7 -1.42	100
BF	10.3 1.46	28.7 7.40	2.3 -3.48	31.0 -0.12	27.6 -1.35	100
SUFF	10.2 1.15	1.7 -1.59	20.3 0.38	6.8 -3.40	61.0 3.15	100
CPX	15.4 2.22	0.0 -1.69	5.1 -1.92	17.9 -1.53	61.5 2.61	100
PREF	0.0 -1.07	0.0 -1.15	61.1 4.26	0.0 -2.39	38.9 0.18	100
CLIP	0.0 -0.67	0.0 -0.71	14.3 -0.25	71.4 1.86	14.3 -0.97	100
Total	6.4	7.3	18.2	31.8	36.3	100

Table 4. Contingency table and Pearson's residuals for Morpho1_V and Origin3_N ($p < 0.05$)

Table 4 shows the results for the pair Morpho1_V and Origin3_N. It confirms the positive correlation between borrowed bases and suffixation as well as complex processes. It also shows a significant negative relationship between affixation and base nouns formed through compounding. Back-formation favors base nouns formed through prefixation and disfavor base nouns which include only one lexeme and no affix. Those bases however favor the formation of verbs through prefixation. Conversion favors bases formed through compounding, but disfavors prefixed base nouns. Finally, complex processes favor base nouns involving suffixation.

4.4. Pair 4: Origin1_N and Morpho2_V

Origin1_N Morpho2_V	EN	B	Total
CONV	68.3 1.07	31.7 -1.42	100
BF_ion	69.8 0.50	30.2 -0.66	100
SUFF_ize	42.5 -1.68	57.5 2.22	100
CPX_ize	44.4 -1.25	55.6 1.66	100
SUFF_OTH	31.6 -1.75	68.4 2.32	100
BF_ing	100.0 1.93	0.0 -2.56	100
BF_OTH	38.5 -1.14	61.5 1.51	100
BF_er	76.9 0.60	23.1 -0.79	100
CPX_SUFF	11.1 -1.98	88.9 2.62	100
PREF_nocomp	62.5 -0.04	37.7 0.06	100
CLIP	85.7	14.3	100

	0.73	-0.97	
PREF_comp	66.7	33.3	100
	0.09	-0.12	
PREF_un	50.0	50.0	100
	-0.34	0.45	
CPX_PREF	66.7	33.3	100
	0.06	-0.09	
Total	63.7	36.3	100

Table 5. Contingency table and Pearson's residuals for Morpho2_V_rec and Origin1_N (p<0.05)

The pair Morpho2_V and Origin1_N confirms the positive correlation between borrowed bases and all types of verbal suffixations. It gives more precision on the correlation between borrowed bases and complex processes as it shows a significant relationship between borrowed bases and complex processes involving suffixation other than *-ize*. Finally, according to the dataset, back-formation in *-ing* appears to be incompatible with borrowed bases. This observation is not surprising considering that *-ing* is a suffix native to English and is very unlikely to be found on borrowings.

4.5. Pair 5: Origin2_N and Morpho2_V

Origin2_N Morpho2_V	Germanic languages	Romance languages	Other languages	Total
CONV	72.2	22.8	5.0	100
	1.05	-2.02	1.12	
BF_ion	69.8	27.9	2.3	100
	0.18	-0.09	-0.50	
SUFF_ize	47.5	52.5	0.0	100
	-1.54	2.82	-1.24	
CPX_ize	48.1	40.7	11.1	100
	-1.22	1.17	1.93	
SUFF_OTH	52.6	47.4	0.0	100
	-0.79	1.52	-0.85	
BF_ing	100.0	0.0	0.0	100
	1.68	-2.27	-0.83	
BF_OTH	38.5	61.5	0.0	100
	-1.27	2.22	-0.71	
BF_er	76.9	23.1	0.0	100
	0.41	-0.38	-0.71	
CPX_SUFF	11.1	88.9	0.0	100
	-2.06	3.38	-0.59	
PREF_nocomp	62.5	37.5	0.0	100
	-0.17	0.47	-0.55	
CLIP	85.7	14.3	0.0	100
	0.59	-0.71	-0.52	
PREF_comp	66.7	33.3	0.0	100
	-0.03	0.21	-0.48	
PREF_un	75.0	25.0	0.0	100
	0.18	-0.14	-0.39	
CPX_PREF	66.7	33.3	0.0	100
	-0.02	0.15	-0.34	
Total	67.5	28.6	3.8	100

Table 6. Contingency table and Pearson's residuals for Morpho2_V_rec and Origin2_N (p<0.05)

The pair Morpho2_V and Origin2_N only shows significant relationships between Romance bases and morphology, with one exception. Romance bases favor suffixation in *-ize*, complex processes involving a suffix other than *-ize*, and back-formation involving the deletion of suffixes other than *-ion*, *-er* or *-ing*. Romance languages disfavor conversion and *ing-* back-formation. As regards the Germanic languages, the data suggests that complex processes involving a suffixation other than *-ize* disfavor Germanic base nouns. No other result is significant for this pair of variables.

4.6. Pair 6: Origin3_N and Morpho2_V

Origin3_N Morpho2_V	MMN_ SUFF	MMN_ PREF	MMN_ SPX	MMN_ COMP	B	Total
CONV	4.1 -1.63	4.1 -2.15	21.3 1.31	38.8 2.29	31.7 -1.42	100
BF_ion	7.0 0.15	41.9 8.39	2.3 -2.44	18.6 -1.53	30.2 -0.66	100
SUFF_ize	10.0 0.90	2.5 -1.12	22.5 0.63	7.5 -2.72	57.5 2.22	100
CPX_ize	18.5 2.49	0.0 -1.40	7.4 -1.32	18.5 -1.22	55.6 1.66	100
SUFF_OTH	10.5 0.71	0.0 -1.18	15.8 -0.25	5.3 -2.05	68.4 2.32	100
BF_ing	5.6 -0.14	33.3 4.09	0.0 -1.81	61.1 2.21	0.0 -2.56	100
BF_OTH	23.1 2.38	0.0 0.97	7.7 -0.89	7.7 -1.54	61.5 1.51	100
BF_er	15.4 1.28	7.7 0.05	0.0 -1.54	53.8 1.41	23.1 -0.79	100
CPX_SUFF	0.0 -0.76	0.0 -0.81	0.0 -1.28	11.1 -1.10	88.9 2.62	100
PREF_nocomp	0.0 -0.71	0.0 -0.76	62.5 2.93	0.0 -1.59	37.5 0.06	100
CLIP	0.0 -0.67	0.0 -0.71	14.3 -0.25	71.4 1.86	14.3 -0.97	100
PREF_comp	0.0 -0.62	0.0 -0.66	66.7 2.78	0.0 -1.38	33.3 -0.12	100
PREF_un	0.0 -0.51	0.0 -0.54	50.0 1.49	0.0 -1.13	50.0 0.45	100
CPX_PREF	33.3 1.85	0.0 -0.47	0.0 -0.74	33.3 0.05	33.3 -0.09	100
Total	6.4	7.3	18.2	31.8	36.3	100

Table 7. Contingency table and Pearson's residuals for Morpho2_V and Origin3_N (p<0.05)

The last pair of variables shows some significant relationships. First of all, as the variable Origin3_N involves the same elements in borrowed bases ("B") as in Origin1_N, the same significant correlations are found here as in section 4.4. The data suggests that conversion favors base nouns formed through compounding and disfavor base nouns including a prefix. Conversely, prefixed base nouns favor back-formation of the *-ion* type. This type of back-formation appears to be incompatible with base nouns involving only one lexical base and no affix. Back-formation of the *-ing* type favors bases formed through prefixation and compounding. Other types of back-formation favor bases

involving a suffix. The data confirms that compounded bases disfavor any type of suffixation. Verbs formed through complex processes involving a suffixation in *-ize* favor suffixed base nouns. Finally, verbs formed through prefixation favor base nouns whose structure involves no affix. The residuals suggest that the nature of the prefix does not impact this result.

5. Discussion and conclusion

After crossing the variables, it appears that some elements are significant when describing the link between the origin of the base noun and the morphology of the verb. However, this relationship is far from being systematic.

In section 2.4, I summarized the expected results of the influence of base noun etymology on the morphology of the derived verb. Some of these results are confirmed by the study. Concerning conversion, the data shows indeed an overrepresentation of Germanic base nouns, either endogenous to English or borrowed from other Germanic languages. For verbs built on endogenous nouns, there is a strong correlation between compound base nouns and conversion. Due to a lack of data, no conclusion can be reached on a potential correlation between nouns borrowed from languages other than Romance or Germanic and conversion. Similarly, the data does not show any significant negative correlation between the process of conversion and suffixed base nouns. However, the data suggests that conversion is underrepresented for bases borrowed from Romance languages as well as endogenous bases formed through prefixation.

As regards verbalization through suffixation, the data confirms that Romance bases are overrepresented. As mentioned in section 3.2.2, most of the results concerning suffixation apply to *-ize* as the data does not include many instances of other suffixes. However, the overrepresentation of borrowed bases and underrepresentation of compound bases are significant results for all cases of suffixation in the data. According to previous studies, it is expected that suffixation never happens with compounded bases. The results suggest otherwise as 4 instances of compounded bases are subject to suffixation: *monomerize_V* < *monomer_N*, *mutagenize_V* < *mutagen_N*, *allotropize_V* < *allotrop_N* and *carboxylate_V* < *carboxyl_N*. However, all these base nouns are formed through the compounding of combining forms. As a result, the data does not contradict the fact that suffixed verbs do not accept compounds as their bases, if compounding is narrowly defined. The four examples in the sample are a consequence of a methodological choice linked to a broad definition of compounding. More surprisingly, the data suggests that suffixation in *-en* does not always attach to Germanic bases contrary to what was expected. There is no statistically significant result for this type of suffixation as the sample only includes one instance: *piecen_V* < *piecen_N*. This example appears to be an exception to the selection rules of the *-en* suffix as the noun *piece* is a borrowing from French. Nevertheless, as this noun has been in the English language since the 13th century (*OED*, 2006), the foreignness of the noun may have been lost when the verb was formed in 1831.

There are not many instances of verbs formed through prefixation in the sample and only two significant results for this process: the overrepresentation of prefixation on simple bases (involving one lexeme and no affix), which was not expected, and the underrepresentation of compounded bases. The lack of sufficient data to carry out a statistical analysis suggests that prefixation is indeed not a productive process in the formation of denominal verbs in English. While a distinction was made in section 3.2.2 between prefixes depending on whether or not they can attach to compounded bases, no instance of a compounded base leading to prefixation is present in the sample. Moreover, it was expected to observe an overrepresentation of suffixed bases for verbs formed through prefixation, but there are no examples of this either.

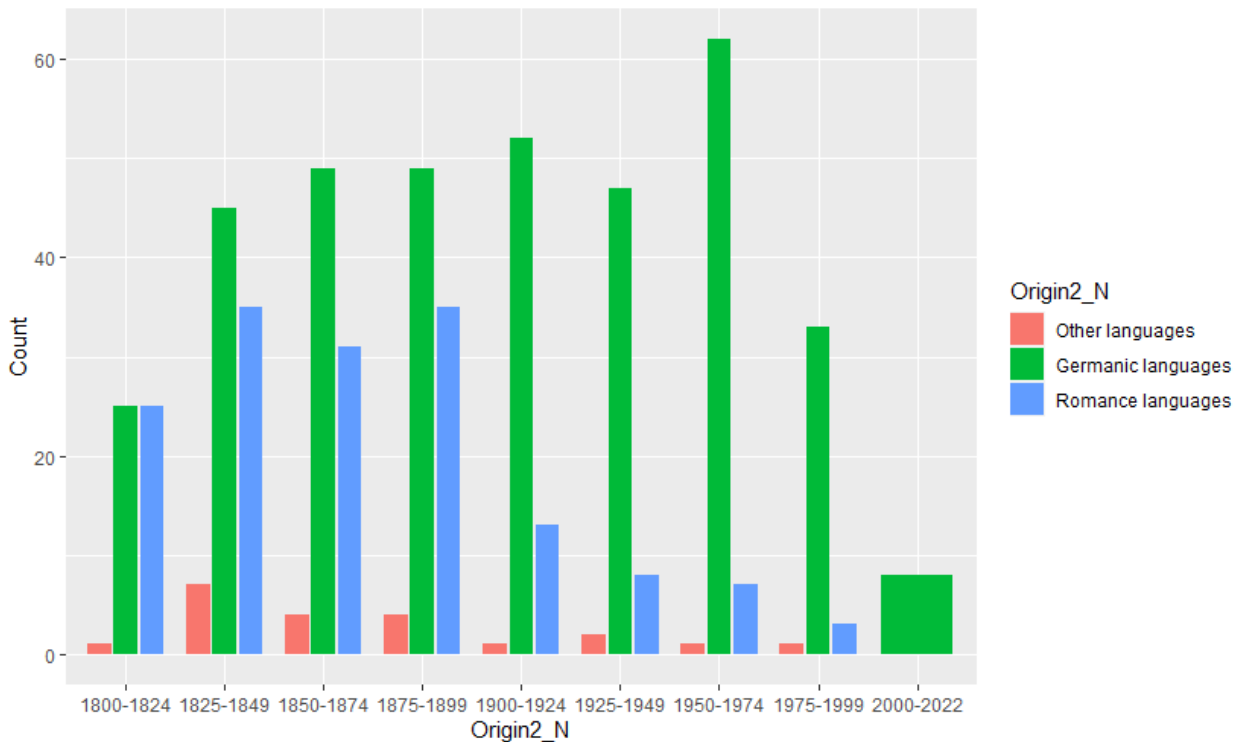
As regards processes which are usually not taken into account in studies on denominal verbs, this study brings preliminary results. They suggest that back-formation favors prefixed bases and disfavors simple bases. The latter result is quite understandable as back-formation involves the deletion of a suffix-like element at the end of the base. To explain the relationship between prefixed

bases and back-formation, it is necessary to take a closer look at the data. Instances of back-formation on a prefixed base all involve words with an ending that can be analyzed as an overt suffix. They are categorized as prefixed bases because the suffixation of the base was anterior to its prefixation. For instance, the verb *self-deport* is built on the noun *self-deportation* through a process of back-formation of the *-ion* type. The base noun is formed through the prefixation of *self-* on the noun *deportation*. This noun was borrowed from Latin but its ending in *-ation* is frequently found in English^{xviii}. The hypotheses I drew for back-formation from previous studies were twofold: the overrepresentation of suffixed bases and of compounded bases. The data suggests that back-formation of the *-ing* favors compounded bases, and disfavors borrowed bases. Other instances of back-formation (excluding *-ion*, *-ing* and *-er*) favor suffixed bases as well as bases borrowed from Romance languages. No significant result was found for back-formation of the *-er* type.

As regards complex processes involving affixation, I expected to observe similar restrictions in these processes as in affixation. Concerning complex processes in general, the data suggests that they favor Romance bases as well as suffixed bases, while they are less likely to be found with Germanic bases. The picture is more nuanced for specific processes: complex processes involving suffixation in *-ize* favor suffixed bases; complex processes involving suffixes other than *-ize* favor Romance bases and disfavor Germanic bases; no significant result was found for complex processes involving prefixation and no suffixation. Consequently, complex processes, which overwhelmingly involve suffixation, show similar restrictions as suffixation: they favor Romance bases and not Germanic bases. No significant result was found for clipping due to very few occurrences.

To conclude, the etymology of the base noun is a relevant parameter for the study of denominal verbs first attested from 1800 to 2022. Both geographical origin and morphological makeup were taken into account, and both yielded some significant results. To the exception of clipping, all word-formation processes show significant affinities or restrictions with certain bases. However, this analysis does not show strict rules but rather tendencies linked to the base noun. Other variables are needed to account for the morphology of the derived verb, not to mention its meaning. As far as the Latinate constraint is concerned, both suffixation and complex processes involving suffixation favor bases borrowed from Romance languages, suggesting that the Latinate constraint is relevant for verbs coined between 1800 and 2022. However, Figure 5 shows the distribution of the verbs over time according to the second origin variable. There is a striking decrease in the use of Romance borrowings for the coinage of new denominal verbs at the beginning of the 20th century. The decrease is confirmed all through the century, with no recorded verb built on a Romance borrowing in the 21st century. Hence, the apparent loss of the Latinate constraint that is pointed out by other studies may only be the consequence of this decrease. A search for borrowings from Romance languages in the *OED* shows that their number has dramatically decreased since the 19th century (15,188 entries between 1800 and 1900; and only 3,488 entries between 1900 and 2022). Further research is needed to account for the probable connection between the decline of the apparent loss of the Latinate constraint and this reduction in Romance loan words^{xix}.

Figure 5. Origin over time (VdenomEN)



Bauer, Laurie. *English word-formation*. Cambridge University Press, 1983.

Bauer, Laurie. "Concatenative Derivation". *The Oxford Handbook of Derivational Morphology*, edited by Rochelle Lieber and Pavol Štekauer, Oxford University Press, 2014, p. 118-35.

Bauer, Laurie, Rochelle Lieber and Ingo Plag. *The Oxford Reference Guide to English Morphology*. Oxford University Press, 2013.

Bleotu, Adina Camelia. *Towards a Theory of Denominals: A Look at Incorporation, Phrasal Spell-Out and Spanning*. Brill, 2019.

Clark, Eve V., and Herbert H. Clark. "When Nouns Surface as Verbs". *Language*, vol. 55, n° 4, 4, 1979, p. 767-811.

Debouzie, Chloé. *Competition between affixation and conversion: Studying noun-to-verb change in Present-Day English*. 2024. Université Lumière Lyon 2.

Dixon, R. M. W. "Deriving Verbs in English". *Language Sciences*, vol. 30, n° 1, 1, 2008, p. 31-52.

Green, Jonathon. *Green's Dictionary of Slang - Digital Edition*. 2020, <https://greensdictofslang.com/>.

Haspelmath, Martin, and Andrea D. Sims. *Understanding morphology*. 2nd ed, Hodder Education, 2010.

Kiparsky, Paul. "Remarks on Denominal Verbs". *Complex Predicates*, edited by Alex Alsina *et al.*, CSLI Publications, 1997, p. 473-499.

Lass, Roger. "Language Periodization and the Concept of 'Middle'". *Placing Middle English in Context*, edited by Irma Taavitsainer *et al.*, Mouton de Gruyter, 2000, p. 7-41.

Marchand, Hans. *The Categories and Types of Present-Day English Word-Formation: a Synchronic-Diachronic Approach*. C.H. Beck, 1969 (2nd edition).

Marchello-Nizia, Christiane, Bernard Combettes, Sophie Prévost et Tobias Scheer (éds). *Grande grammaire historique du français*. De Gruyter Mouton, 2020.

Nevalainen, Terttu, and Elizabeth C. Traugott (eds). *The Oxford Handbook of the History of English*. Oxford University Press, 2012.

Oxford University Press, ed. *Oxford English Dictionary [online]*. 2022, www.oed.com.

Plag, Ingo. *Morphological productivity: structural constraints in English derivation*. Mouton de Gruyter, 1999.

Romaine, Suzanne. *The Cambridge History of the English Language. Volume 4: 1776-1997*. Cambridge University Press, 1998.

[RStudio] Barnier, J., F. Briatte and J. Larmarange. *_questionr: Functions to Make Surveys Processing Easier_*. R package version 0.7.7, 2022, <https://CRAN.R-project.org/package=questionr>.

[RStudio] Meyer F. and V. Perrier. *_esquisse: Explore and Visualize Your Data Interactively_*. R package version 1.1.1, 2022, <https://CRAN.R-project.org/package=esquisse>.

[RStudio] Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.

[RStudio] Wickham H. *forcats: Tools for Working with Categorical Variables (Factors)*. R package version 1.0.0, 2023, <https://CRAN.R-project.org/package=forcats>.

[RStudio] Wickham H. *et al.*, “Welcome to the tidyverse”. *Journal of Open Source Software*, 4(43), p. 1686, 2019.

[RStudio] Wickham H. and J. Bryan. *_readxl: Read Excel Files_*. R package version 1.4.0, 2022, <https://CRAN.R-project.org/package=readxl>.

ⁱ *Overrepresentation* is defined statistically as the representation of a group in a category that goes beyond what could be expected if two variables were independent (null hypothesis). Independent variables are variables which do not influence one another, there are not correlated. Let us hypothesize that the origin of the base noun and the morphology of the denominal verb are two independent variables. If 30% of the base nouns are from Romance origin, then we would expect each type of verbal morphology to include 30% of verbs originating from Romance base nouns (30% for conversion, suffixation, etc.). If the proportion is significantly higher, Romance verbs are said to be overrepresented in that category, if it is lower, they are underrepresented.

ⁱⁱ The analysis in terms of parasynthesis is based on the etymology section of the *OED*. Among the few instances present in the data, there are some problematic cases where an alternative etymology could be argued. For instance, the verb *disinsectize* could as well be a back-formation from the noun *disinsectization* which was first attested a decade earlier.

ⁱⁱⁱ The dataset used for this study is available on <https://nextcloud.univ-lyon3.fr/s/L82LcqrzepyHelg>.

^{iv} It is always hard to define relevant temporal boundaries as any choice will draw an artificial line where a continuum should be: “periodization is always conventional, artificial, and subject to different interests, questions, and methods” (Nevalainen & Traugott, 2012: 3, quoting Lass, 2000). The boundaries chosen for the study are based on *The Cambridge History of the English Language*, and more specifically on the fourth volume entitled 1776-1997 (Romaine, 2008). The re-interpretation of the boundaries to 1800-2022 is dictated first by the broader context of the research project which aims at contrasting the English data with French data (see Marchello-Nizia *et al.*, 2020, for the periodization of French) and secondly, by the will to take the most recent period into account.

^v The dates mentioned for lexicographic resources are of two kinds, according to the data available: for *OED* entries, the date corresponds to the last update of the entry; for the *GDS*, the date is the date of collection.

^{vi} Bluegrass, v.: “(US drugs) to commit a drug user (or oneself) to the Lexington Federal Narcotics Hospital in Lexington, Kentucky” < Bluegrass State, PN: “the nickname of Kentucky” (*GDS*, 2022).

^{vii} Napoleonize, v.: “to occupy or govern (a country) in the manner of Napoleon I” < Napoleon, PN: “the name of Napoléon Bonaparte” (*OED*, 2003).

^{viii} As a result of this methodological choice, endogenous base nouns can include nouns involving stems which were borrowed from languages other than English. For instance, the verb *revolver* is formed through the conversion of the noun *revolver*, itself formed through the suffixation of the verb *revolve*, which in turn was

borrowed from both French and Latin (*OED*, 2010). As the noun *revolver* is built through suffixation, it follows a word-formation process within English and is thus categorized as endogenous.

^{ix} People-watch, v.: “to pass time in a public place by idly observing people” < people-watching, n.: “idle observation of people in a public place, as a way of passing time.” (*OED*, 2005).

^x Ritualize, v.: “to practise ritualism” < ritual, n.: “the prescribed form or order of religious or ceremonial rites” (*OED*, 2010).

^{xi} Password, v.: “to make (one's way) by giving a password” < password, n.: “a selected word or phrase securing admission, recognition, etc., when used by those to whom it is disclosed” (*OED*, 2005).

^{xii} Greatcoat, v.: “to dress in a greatcoat” < greatcoat, n.: “a long heavy overcoat” (*OED*, 2013).

^{xiii} Reviver; v.: “to treat (clothing) with a reviver; to renovate” < reviver, n.: “Something which has a reviving, restorative, or invigorating effect” (*OED*, 2010).

^{xiv} Metasomatize, v.: “to alter (a rock, mineral, etc.) by metasomatic processes” < metasomatism, n.: “change in the chemical composition of a mineral or rock, esp. by the dissolution of one mineral and deposition of another” (*OED*, 2001).

^{xv} Non-intervene, v.: “to refrain from intervening. Usually ironically: to intervene” < non-intervention, n.: “absence of intervention; spec. the principle or practice of not becoming involved in others' affairs, esp. as adopted by one state in regard to others” (*OED*, 2010).

^{xvi} Boob, v.: “to treat (a person) as a fool; to deceive or dupe; to trick (a person) into doing something, or out of a possession” < boob, n.: “a stupid, inept, or blundering person; a fool” (*OED*, 2018).

^{xvii} Weaponize, v.: “to adapt for use as a weapon; spec. to provide (a nuclear or other explosive device) with a mechanism for being launched and propelled toward a target” < weapon, n.: “an instrument of any kind used in warfare or in combat to attack and overcome an enemy” (*OED*, 2008).

^{xviii} A search of “*ation” in the *OED* yields 8,335 nouns.

^{xix} I would like to thank the final reviewer for this suggestion as well as their detailed reviewing of the article.