Salvador Valera*

The semantics of noun-to-verb zero-derivation in English and Spanish

Abstract: The semantics of zero-derivation/conversion has attracted renewed interest both as a subject of description and as a means towards refined descriptions of the process. This paper takes the latter stance and compares which semantic categories occur in zero-derivation/conversion and in overt affixation in two languages with a different morphological model: English and Spanish. For attestation and distribution of the semantic categories, the paper relies on a stratified sample of denominal verbs collected from two comparable corpora: the *British National Corpus* and the *Corpus de Referencia del Español Actual*. The sample consists in sets of ca. 50 denominal verb-forming resources, one per affixation process, namely affixation by *-ate*, *-(i)fy*, *-ize/-ise*, *-en*, *en-* in English, *a-…-ar*, *en-…-ar*, *-ear*, *-ecer*, *-(i)ficar*, *-izar* in Spanish, and zero-derivation/conversion in both languages. The results are contrasted within and across affixes/processes, within each language and across the two languages for the adequacy of a description as zero-derivation or as conversion. Statistical analysis shows that the process/affixes form a cline with overlaps and closer associations between specific affixes/processes and semantic categories, but no clear divide between zero-derivation/conversion and the rest of processes as far as the semantic categories are concerned.

Keywords: zero-derivation, conversion, affixation, semantic categories, English, Spanish

1 Introduction

The semantic range of words formed by zero-derivation/conversion has become a renewed field of interest, as can be seen from the interest raised by the workshop that gave rise to this paper as well as from recent publications on the subject (e. g. Baeskow 2020, among others). The semantic range of zero-derivation is relevant in itself as regards what semantic categories can be attested by corpus data. Research on such semantic range also offers the possibility of contrasting the results

*Corresponding author: Salvador Valera,** Department of English and German Philology, University of Granada, Granada, Spain, e-mail: svalera@ugr.es

obtained with those of the semantic range of comparable formations. Ultimately, the objective is to assess to what extent noun-to-verb zero-derivation differs from denominal derivation of verbs by overt affixes as regards their semantic ranges, i. e. how (dis-)similar zero-derivation is from overt affixation regarding the semantic patterns found in denominal verb formation and, thus, whether a case for a zero affix can be sustained on these grounds, or not.

A preliminary remark is in order with regard to this paper: the approach presented here views zero-derivation as a dynamic, asymmetrical process, i. e. as a derivational process whereby lexemes of the same form can be classified as different word-classes (cf. Bauer and Valera 2005: 8 for English and Pena 1991: 103 for Spanish,[1] among others) and, in the case of denominal verb formation, that the change in the word-class involves a specification of the contrast from ENTITY to PROCESS/ACTION (Štekauer 1996):

(1)   *bottle*   vs.   *to bottle*
(2)   *agua*   vs.   *aguar*
       'water'   'to water'

This article relies on statistical analysis of corpus data for a comparison of whether zero-derivation follows or deviates from the patterns presented from denominal verb formation by overt affixation. To that end, it reviews the literature with regard to zero-derivation in English and Spanish (Section 2), presents the method of the empirical research presented here (Section 3) and discusses the implications of the results obtained (Section 4).

2 Zero-derivation in English and Spanish

In general, zero-derivation relies on the conditions of word-class change and formal identity (cf. Tournier 2010 [1985]: 171, among others; Valera 2015 and cf. e. g. note 2).

The main point to be raised as regards theoretical principles concerns two opposite views in the literature of Spanish besides the hypothesis of a derivational zero (cf. Valera 2021). Thus, in Spanish morphology two main positions can be found:

i)   affixation, substitution and subtraction are maximized at the cost of conversion, as in the tradition of Romance linguistics (e. g. as in Pena 1991; 1999), or

---

1 Unlike English, this description for Spanish allows the contrast to involve word sub-class instead of word-class. The theoretical possibility, described in English as secondary word-class conversion (Leech 1974: 214–216), has rarely been echoed in the literature.

ii) conversion is maximized at the cost of affixation, substitution and subtraction, as in the tradition of Germanic linguistics (e. g. as in Rainer 1993; 2016).

In addition to these two, an alternative (and less frequent) approach is available as zero-derivation, whereby a zero affix takes the position after the verb's thematic vowel to operate the word-(sub)class change (cf. Varela Ortega 1990: 95; Pena 1991: 107–109, 2012).

The well-known advantages and disadvantages of each approach apply similarly in English and Spanish (in the latter, also in the maximized approach to affixation): zero-derivation explains unmarked word-class change with a simpler system, but it does at the cost of a number of rules (for separate formations by the same allomorph) and maximized homophony (cf. Dahl and Fábregas 2017; Harley 2020, among others); conversion accounts for unmarked word-class change, but it does at the cost of a more complex framework with two different structure-building systems (affixation and conversion).[2] Advocates of zero-derivation argue that a separate process can be avoided, if a derivational counterpart to inflectional zeroes is posited. Advocates of conversion view derivational zeroes as unnecessary, artificial descriptive devices.

For want of other evidence and in view of the theoretical deadlock blocking the resolution of the debate, this paper explores the semantic range of denominal verb formation in English and Spanish. Assuming that unmarked noun-to-verb change may be by a zero affix, this paper compares it with its overt affixal counterparts for similarities and differences. In the long run, and according to what other word-classes and directionalities may reveal as regards the semantic categories involved, this type of evidence may cast light on whether the process supports a theory of derivational zeroes, or a theory of conversion.

3 Method

3.1 The sample

This paper relies on data-based statistical analysis of two comparable sets of verbs formally identical to nouns. The sets consist in groups of ca. 50 types extracted

---

2 As pointed out by an anonymous referee, conversion does not necessarily entail additional complexity to the descriptive framework: some theories, e. g. Becker (1990), present word-formation rules as triples of operations mapping syntactic category, form, and semantics [such that in zero-derivation, unlike suffixation] there is no difference between input and output in the formal part of the rule [so] no two structure building frameworks are needed'.

Table 1: Data subsample profile according to query for English and Spanish.

	English	Spanish
Prefixation	*em-/en-*	*a-…-ar*
		en-…-ar
Suffixation	*-ate*	*-ear*
	-en[3]	*-ecer*
	-(i)fy	*-(i)ficar*
	-ize	*-izar*
Zero-derivation	[Tags _N, _V]	*-ar* [with related N form]

over a range of frequencies from two ca. 100 million-word comparable corpora: the *British National Corpus* for English (hereafter, BNC) and the *Corpus de Referencia del Español Actual* for Spanish (hereafter, CREA). The aim is not so much to try and find a unified account of zero-derivation, but to research its semantic range in two such morphologically different languages.

The sets can be divided into two subsamples (one per language) of productive denominal verbalizing affixation yielding various semantic categories, as in Table 1 (Plag 2018 [2003]; Serrano-Dolader 1999).

A number of methodological provisos were in effect for use of the sample, the most relevant being:

i) *-ar*: Only verbs in *-ar* were sampled for the Spanish dataset, because the other two types, *-er* and *-ir* verbs, are considered no longer productive according to both the literature (cf., among others, Serrano-Dolader 1999) and the samples collected.[4]

ii) Bases are nouns: The verbs that could be traced back both to nouns but also to formally identical adjectives were discarded and replaced by a verb formed by the same process where the base can be only a noun. Replacement was within the same or closest possible frequency range as the original type sampled, e. g. *chino* 'Chinese' for *achinar* 'make look like Chinese'.

The data sample consists in ca. 50 types for each of the entries listed in Table 1. The types were selected at random starting out from frequency range 1[5] to en-

---

3 As pointed out by the editors, Plag (1999) considers *-en* suffixation no longer productive.

4 This, inferred from the fact that only one type was recorded as the result of the queries for the (*a-*)-*er*/-*ir* and (*en-*)-*er*/-*ir* groups, namely *enlucir* (*en-luz-ir*) 'plaster'.

5 The frequency ranges used for sampling are according to the corpora frequency lists available, not to the concordances actually retrievable from the corpora. For this sample, the difference ranges between −490 types (for zero-derivation) and −15 types (for *-ize* affixation) retrievable

Table 2a: Sample types of the English sample classified by process/affix after replacement of irrelevant cases.

en-	*-ate*	*-en*	*-(i)fy*	*-ize/-ise*	Ø
48	48	11	23	44	49

Table 2b: Sample types of the Spanish sample classified by process/affix after replacement of irrelevant cases.

a-…-ar	*en-…-ar*	*-ear*	*-ecer*	*-(i)ficar*	*-izar*	Ø
50	50	50	5	50	50	50

sure that productive patterns are represented in the sample as far as possible. Frequency 1 was therefore pursued, following Baayen and Lieber (1991), on the principle that hapaxes and so-called pseudo-hapaxes signal productive formations. Types selected originally at random but of no value (e. g. names, misspellings and similar cases) were replaced within or as close as possible to the frequency range of the type selected originally at random. When no new replacements were possible, supplementary frequency ranges were used, but not to the extent of going over the highest frequency range of the set of types of the process. As a result, when no relevant replacements were available, the set of types for specific processes or affixes falls below the target 50-type size, as detailed in Tables 2a and 2b.

The types are listed in Tables A1 and A2 for English and Spanish respectively. The sample takes account of the polysemanticity of each of these processes/affixes by recording the semantic category attested in each token of the types in the sample. This means that every sense of every type (i. e. the semantic category it formalizes) is recorded. The resulting sample is highly representative of the types researched, in that it captures the entire semantic range of each type by the analysis of the senses attested in the concordances retrieved from the BNC and from the CREA. Overall, the results rely on evidence based on 17988 tokens for English and 11083 for Spanish, as listed in Tables 3a and 3b.

---

from the corpus compared with the BNC frequency list, and between −5 types (for *-(i)ficar* and +54 (for *a-…-ar*) retrievable from the CREA corpus.

Table 3a: Number of tokens classified by process/affix, according to their BNC concordances.

en-	-ate	-en	-(i)fy	-ize/-ise	Ø
4344	3877	1391	2489	2240	3657

Table 3b: Number of tokens classified by process/affix, according to their CREA concordances.

a-…-ar	en-…-ar	-ear	-ecer	-(i)ficar	-izar	Ø
2607	1987	1256	11	1660	1391	2171

3.2 Semantic classification

The semantic categories considered for the dataset are based on categories described in the literature (for English, Marchand 1969 [1963]: 365; Kastovsky 1994: 97–98; Clark and Clark 1979; Karius 1985; Quirk et al. 1985: 1560; Plag 1999: 219–220; Štekauer 1996; Tournier 2010 [1985]: 185; Cetnarowska 1993: 86; and references cited therein; for Spanish, Alemany Bolufer 1920: 139–139; Lázaro Carreter 1980: 75; Rainer 1993: 213–215, 237–239; Serrano-Dolader 1999; Fábregas 2016; Santiago Lacuesta and Bustos Gisbert 1999: 4514–4518, 4550), and on corpus data attestation, as shown in Table 4.

Table 4: Semantic categories considered for the dataset. Unless otherwise specified, examples are from the BNC and the CREA, and glosses are from the *Oxford English Dictionary* and *Diccionario de la Real Academia Española*.

Based on references	English	Spanish
LOCATIVE	to jail[6] 'to confine in or as in a jail; to imprison'	*estacionar*[7] '*situar en un lugar, colocar*' 'to put in a location, to place'

---

6 *He launched his campaign on Saturday with a promise to protect the public sector and to jail those caught stealing from the state.*

7 *La Alianza, pese a que no puede estacionar tropas en los tres nuevos países de forma permanente, tal y como acordó con Rusia, contará con nuevos campos de maniobras en lugares de mucha menor densidad de población que los occidentales, como Polonia.*

Table 4 (continued)

Based on references	English	Spanish
ORNATIVE	to glass[8] 'to fit or fill in with glass'	*premiar*[9] '*otorgar un premio a alguien*' 'to give an award to somebody'
CAUSATIVE	to jelly[10] 'to 'set' as jelly; to congeal, solidify, coagulate'	*eficienciar*[11] '*hacer eficiente*' [my gloss] 'make efficient'
RESULTATIVE	to hook[12] 'to make hook-like or hooked; to bend, crook, incurve'	*parcelar*[13] '*medir, señalar las parcelas para el catastro*' 'to measure, to mark plots for the land registry'
INCHOATIVE	to widow[14] 'to become the widow of'	
PERFORMATIVE	to blare[15] to roar with prolonged sound in 'weeping, as a child; to bellow as a calf. [...] to sound loudly and stridently'	*contactar*[16] '*establecer contacto o comunicación con alguien*' 'to establish contact or communication with somebody'

---

8 *There are two chic consulting areas decorated in cream and maroon, glassed off for privacy.*

9 *La filosofía de esta medida es premiar a los empresarios verdes y castigar a los que ensucien el ambiente sin dañar por ello a los consumidores ni perjudicar a los costes de las pequeñas y medianas empresas.*

10 *She took all the bones out of his spinal column and compressed them to nothing, she turned his eyeballs inward, she jellied his legs, she unplugged and rewired his guts.*

11 *Dijo que conjuntamente con el nuevo director de la Policía Escolar, general Luis Reynoso López, va a diseñar un plan para aumentar la seguridad de los centros educativos y eficienciar el uso de la Policía Escolar.*

12 *You may hook your toes under a chair if you like, but do not link your fingers behind your head as this puts strain on your neck.*

13 *Está prohibido parcelar o dividir un inmueble con el fin de venderlo o arrendarlo por lotes, según los decretos de creación de los distintos distritos de riego y la Ley de Riego y Avenamiento.*

14 *In August 1561 Mary, widowed after a short marriage and a short reign [...] returned to her own kingdom, landing in the mist at Leith and making her way to Holyroodhouse.*

15 *There by the window was the portable gramophone all ready to blare out the Hrst* [sic] *Wessel the very next morning.*

16 *Querría contactar con algunas de las personas que ya hayan recibido el cheque vivienda para poder informarme sobre cómo y dónde intentar pedir este tipo de ayudas.*

Table 4 (continued)

Based on references	English	Spanish
SIMILATIVE	to bullock[17] 'to work like a bullock, i. e. strenuously without intermission'	*reinar*[18] '*dicho de una cosa: prevalecer o persistir continuándose o extendiéndose*' 'said of a thing: to prevail or persist by continuing or extending itself'
INSTRUMENTAL	to decoy[19] 'to alure or entice (wildfowl or other animals) into a snare or place of capture: said usually when this is done by, or with the aid of, another animal trained to the work'	*cincelar*[20] '*labrar con cincel piedras o metales*' 'to chisel stones or metals'
PRIVATIVE		
STATIVE		*incordiar*[21] '*molestar, importunar*' [*ser un incordio (para alguien), i. e. 'persona o cosa incómoda, agobiante o muy molesta*'] 'to annoy, to intrude' [to be a nuisance (to somebody)]
Corpus attested-only		
EFFECTIVE	to receipt[22] 'to give a receipt for (a sum of money, purchase, etc.)'	*fabular*[23] '*inventar cosas fabulosas*' 'to invent fabulous things'

17 *Just to complete the amazing change in fortunes, Mike Allingham ripped on to a short lineout throw from Kenny Milne, who then took his scrum-half's pass and bullocked his way over from about ten metres out, Lawrie adding the two points.*
18 *Esta es, a mi juicio, la función más propia de un rey. En ello consiste eso que se llama reinar.*
19 *In which case they would try to decoy the enemy towards some of the hidden groups, where they would be outnumbered and beaten.*
20 *A justificar su clase otra vez, como cuando tuvo que cincelar su cuerpo y engordar más de diez kilos en músculo para dar el definitivo salto a los pesados, donde se mueven las grandes bolsas y donde se adquieren las entradas para instalarse en la leyenda.*
21 *Desde que Schroeder fue derrotado por Scharping en la elección interna de 1993 para elegir al nuevo líder socialdemócrata, el primero no ha dejado de incordiar al segundo.*
22 *In the absence of a printed form, a letter giving details of the charge is sufficient. However, always send a copy for receipting and return.*
23 *Soy consciente de que la memoria no es fidedigna porque es muy fácil, y sobre todo para los escritores, dejarse llevar, manipular el recuerdo y fabular sobre él.*

Table 4 (continued)

Based on references	English	Spanish
DIRECTIONAL	to nightclub[24] 'to visit or go to a nightclub'	
OTHER	to weather[25] 'to withstand and come safely through (a storm); to come safely through (a period of trouble, adversity, affliction, etc.); to sustain without disaster'	*estiar*[26] '*alcanzar el estío*' [my gloss] 'reach the [state of] summer'

Each token was classified manually according to how the definitions of each type in the *Oxford English Dictionary* (hereafter, OED) and in the *Diccionario de la Real Academia Española* (hereafter, DRAE), match the paraphrases of the semantic categories described for denominal verbs formed by zero-derivation. This entailed both complete matches between the sense attested by concordances and the OED definitions (e. g. *melodramatise* defined in the OED as 'to make melodramatic', and therefore classified as CAUSATIVE), but also incomplete matches (e. g. *exemplify* defined in the OED as 'to support, illustrate, or demonstrate (an assertion, general rule, etc.) with an example or examples'). Incomplete matches were classified as semantic categories by approximation of the hyponym/hypernym in the definition (as ORNATIVE in the former example with *exemplify*). Consultation with two additional researchers was held whenever semantic classification was unclear, e. g. in order to decide on one semantic category per concordance, if several are possible (e. g. as in *codify* defined as 'To reduce to a general system; to systematize', hence classifiable as ORNATIVE but also as RESULTATIVE).[27]

A major point in the study is the directionality of the process. While the focus is on denominal verbs, no assumption of directionality was made; random sampling included items regardless of any presumed directionality. Noun-to-verb

---

24 *Ordinarily, there would be very few behaviour problems with the dogs of Barbados, even if we were viewing the scene with the rose-tinted spectacles of tourists arrived for two weeks of beaching, sightseeing and nightclubbing!*

25 *In this compelling new video, Peters offers six ways to weather a recession and shows us how we can use recession as an opportunity.*

26 *Cuando la República también dejaron estiar las fincas, por no colaborar.*

27 In case of disagreement, internal coherence within the sample was used as an additional criterion, such that, if one category is more frequent than another and both may apply in one concordance, the more frequent one is primed unless it went against a majority of three analyses.

directionality was thus confirmed by way of the match with the semantic categories listed in the literature for denominal verb formation. If the concordances of a type did not follow the semantic categories expected from noun-to-verb formation, analysis as deverbal zero-derivation into nouns was attempted, to cater for the opposite directionality following Sanders (1988) and Bauer and Huddleston (2002) for English, and Alemany Bolufer (1920) and Alvar Ezquerra 2015 [1994] for Spanish. If this operation of reversed verb-to-noun paraphrasis was successful, then verb-to-noun directionality was considered to be attested and the type was deleted from the dataset. If interpretation alongside the patterns for deverbal noun formation still did not fit, reinterpretation was attempted following the semantic categories listed in Bagasheva (2017) for cross-linguistic research, whether they entailed noun-to-verb or verb-to-noun zero-derivation (e. g. as in *to kitten* meaning 'give birth to kittens', EFFECTIVE). At this stage, noun-to-verb formations were retained and verb-to-noun formations were deleted from the dataset.

It is worth recalling at this point that, to the best of my knowledge, the references on the range of semantic categories that can be expressed by zero-derivation do not specify the proportion to which each category occurs, whether in English or in Spanish. The default hypothesis is that not all occur to the same degree, i. e. not all are equally possible, just as not all the semantic categories expressed by overt affixes occur to the same degree.

3.3 Statistics

The data obtained from the semantic classification of the senses recorded as detailed in Sections 3.1 and 3.2 produce a complex dataset for which ordination analysis is an adequate method.

In order to examine whether some variables (processes and number of concordances within each process for a semantic category) can explain the similarity between words regarding their different representation on the various categories, Canonical Analysis (hereafter, CA) is used for the presence/absence of semantic categories in the different words, and Partial Redundancy Analysis (hereafter, RDA) is used for the frequency of semantic categories in the types, i. e. for abundance.

For the quantitative study with RDA, the frequency of occurrence of each category per word was expressed as a percentage with respect to the total number of concordances attested for each category per word. In CA and partial RDA analyses, *Processes* and *Number of concordances* were included as constraining variables. Forward selection was applied to the overall models to increase parsimony and reduce correlation between explanatory variables. ANOVA-like permutation tests

to assess the significance of the models and constraining variables were employed after 999 permutations.

The dataset was thus subject to statistical analysis:

i)   CA considers several datasets at the same time and examines their interrelations. CA is used here for the identification of the presence or absence of semantic categories in the samples of types that make the entire dataset;

ii) Principal Component Analysis (hereafter, PCA) considers datasets for the identification of variables to explain variation within a sample of simplified results; and

iii) RDA considers datasets for the identification of the best possible explanatory variable, removing the effect of other possible explanatory variable(s). RDA is used here for a quantitative study of the frequency of the semantic categories in the sample of types, i. e. the abundance of the semantic categories across the types in each language.

The analyses examine the effect of the explanatory variables as described above. In order to identify as many associations as possible, unconstrained ordination analysis of response variables was included to explore their relationship without constrains (Borcard et al. 2011). Therefore, unconstrained CA (for presence/absence) and PCA (for abundance) were also run.[28] All analyses and plotting were conducted with R software 3.6.1 (R Core Team 2019), using the packages vegan (Oksanen et al. 2019), PCNM (Legendre et al. 2012) and MASS (Venables and Ripley 2002).

4 Results

This section presents the results according to the statistical analysis cited in Section 3.3. Several graphic representations may be possible, e. g. in the case of three-dimensional representation of PCA. In what follows only the clearest representations are shown.

---

28 PCA is intended to reduce the dimensions of large datasets, as the ones used in this paper, and reveal the clearest patterns contained therein. PCA has been cited for similar research since as early as Baayen (1994), and has been used for specific language research topics (e. g. as in Gries 2003 or, more recently, in Savoy 2012 or Carling et al. 2013 for similar datasets/purpose as the ones used here).

4.1 Canonical analysis

For English, unconstrained CA after removal of the semantic category PRIVATIVE (not attested in the sample) and of types that have not been classified under any of the semantic categories described in the literature, whether for English or for Spanish (whereby the total number of types researched decreases from 223 to 217) shows that, according to Kaiser-Guttman, eight axes can explain between 7.8 % and 9.6 % of the total data variance. This means that the occurrence of types according to their semantic categories can hardly be explained thus or, in other words, the types do not show a clear pattern as regards their occurrence according to semantic category and is therefore inconclusive. Thus, type ordination according to the presence/absence of each of the semantic categories does not lead to firm conclusions.

For Spanish, CA after removal of the categories PRIVATIVE and PLEONASTIC (not attested in the sample) and of types that have not been classified under any of the semantic categories described in the literature, whether for English or for Spanish (whereby the total number of types decreases from 304 to 303) shows that, according to Kaiser-Guttman, eight axes can explain between 8.7 and 9.9 % of the total data variance, i. e. moderately higher than in English. This means that, like in English, the occurrence of entries according to their semantic categories can hardly be explained thus or, in other words, the entries do not show a clear pattern of occurrence according to semantic category and, therefore, semantic category is inconclusive.

4.2 Principal component analysis

PCA is carried out after deletion of the semantic categories PRIVATIVE for English and PRIVATIVE and PLEONASTIC for Spanish (because not any is recorded for any token of any type) and after checking that all types are classified at least as one semantic category. For PCA, the percentage of occurrence of each semantic category per type is calculated. This is intended to standardize the weight of each semantic category according to their frequency of occurrence.[29] The similarities of each type are then compared according to the percentages of each semantic category. The data are analyzed according to which type shows a similar/different distribution of their concordances with regard to the semantic categories. The results are

---

29 Normalization for PCA did not have undesirable effects as regards, e. g. the weight of a rare category in a rare type. The rare categories appear in frequencies above hapaxes or quasi-hapaxes in the two languages.
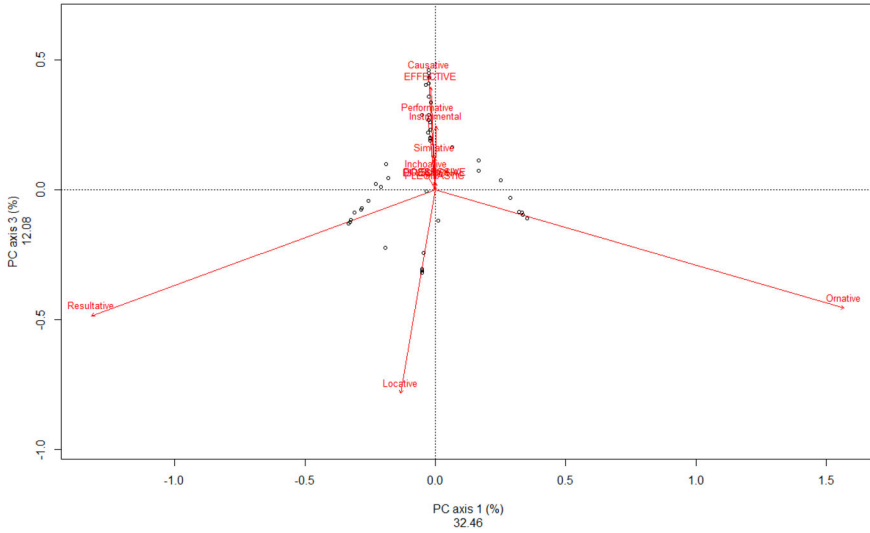
Figure 1: A three-dimensional representation of PCA for the English data sample.

represented in Figures 1 and 2 such that the closest points stand for types with a similar composition (by percentage) of semantic categories within their semantic ranges, and the points that are farther away stand for types with a different composition (by percentage) of semantic categories within their semantic range. The representation shows the axes that can best explain data variation.

For English, PCA describes the data matrix as in Figure 1. Figure 1 is a graphical representation of the explanatory value of each semantic category, as signaled by their position with respect to the rest of categories. Variation is thus revealed, again following Kaiser-Guttman, in terms of three axes that account for 32.5%, 22.7% and 12.2% of the data as follows:

i) Principal Component (PC) 1 explains the highest percentage of the data and separates the categories RESULTATIVE form ORNATIVE.

ii) PC2 explains the intermediate percentage of the data and separates RESULTATIVE and ORNATIVE from LOCATIVE.

iii) PC3 explains the lowest percentage of the data and separates RESULTATIVE, ORNATIVE and LOCATIVE from all other categories, esp. from CAUSATIVE and EFFECTIVE.

Thus, the categories that explain most of the variation are RESULTATIVE, ORNATIVE and LOCATIVE and, to a lesser extent, CAUSATIVE and EFFECTIVE.
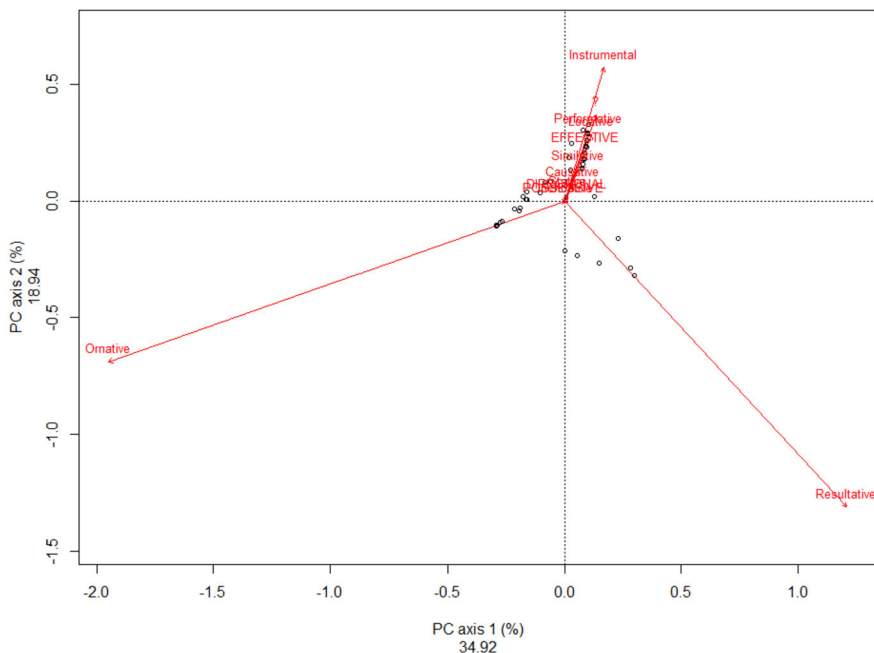
Figure 2: A three-dimensional representation of PCA for the Spanish data sample.

For Spanish, PCA describes the data matrix as in Figure 2. Figure 2 is a graphical representation of the explanatory value of each semantic category, as signaled by their position with respect to the rest of categories. Variation is thus revealed, following Kaiser-Guttman, in terms of four axes that account for 34.9 %, 18.9 %, 10.4 % and 8.8 % of the data. For easier comparison with the English data, only the three most explanatory axes are described below and represented in Figure 2:

i) PC1 explains the highest percentage of the data and separates the categories RESULTATIVE form ORNATIVE, just as in English.

ii) PC2 explains the first intermediate percentage of the data and separates RESULTATIVE and ORNATIVE from INSTRUMENTAL.

iii) PC3 explains the second intermediate percentage of the data and separates INSTRUMENTAL from all other categories.

Thus, the categories that explain most of the variation are RESULTATIVE, ORNATIVE and INSTRUMENTAL.

4.3 Partial redundancy analysis

Based on the above data description, RDA is applied to find the best possible explanation for the data matrix according to the variable *Affix/Word-formation process*. The objective is to identify whether the affix/word-formation process and the number of cases may explain the abundance pattern of each semantic category.

For English, RDA reveals a significant model such that the variables *Word-formation processes/Affix* and *Number of types* explain the variation of similarities between types (F = 5.43, Df = 6–210, p-value = 0.001). Axes RD1 and RD2 explain the variation significantly (RDA 1: Variation explained = 9.1 %, F = 22.02, Df = 1–210, p-value = 0.001; RDA 2: Variation explained = 2.8 %, F = 6.76, Df = 1–210, p-value = 0.003).

The variable *Number of types* does not yield a significant explanation and is discarded after forward selection. The final model after variable selection therefore relies only on the variable *Word-formation processes/Affix*, whereby the model remains significant (F = 6.49, Df = 5–211, p-value = 0.001). Only the variable *Word-formation processes/Affix* explains similarities between types significantly (F = 6.47, Df = 5–210, p-value = 0.001).

RDA of the variable *Word-formation processes/Affix* accounts for 13.34 % of the variation by two axes, represented in Figure 3 (RDA 1: Variation explained = 9.05 %, F = 22.03, Df = 1–211, p-value = 0.001; RDA 2: Variation explained = 2.8 %, F = 6.70, Df = 1–211, p-value = 0.003).

Based on the above, the following claims can be made for English:

i) All *Word-formation processes/Affixes* except *en-* prefixation arrange themselves on a continuum with partial overlap as regards their representation of semantic categories. The overlaps vary according to each case, e. g. *-ize/-ise* and *-ify* overlap considerably, whereas *-ate* and zero-derivation overlap considerably less and zero-derivation hardly overlaps with any other of the suffixes (cf. only *-ate*).

Suffixes cluster along a vertical band with *-ify* and *-en* suffixation at each end, and zero-derivation alongside the other affixes in between both ends, even if it lies closer to the end of *-en* suffixation.

ii) Different types of affixation such as prefixation and suffixation appear separate, in that *en-* prefixation lies farthest from all other groups. This is so, even if *-en* suffixation lies relatively separate from all other suffixes too (not as much as *en-* prefixation, though).

iii) Specific associations between affixes and categories appear, e. g.:

a. *en-* prefixation and LOCATIVE appear associated, in line with the literature, attested in 62 % of the *en-* types.
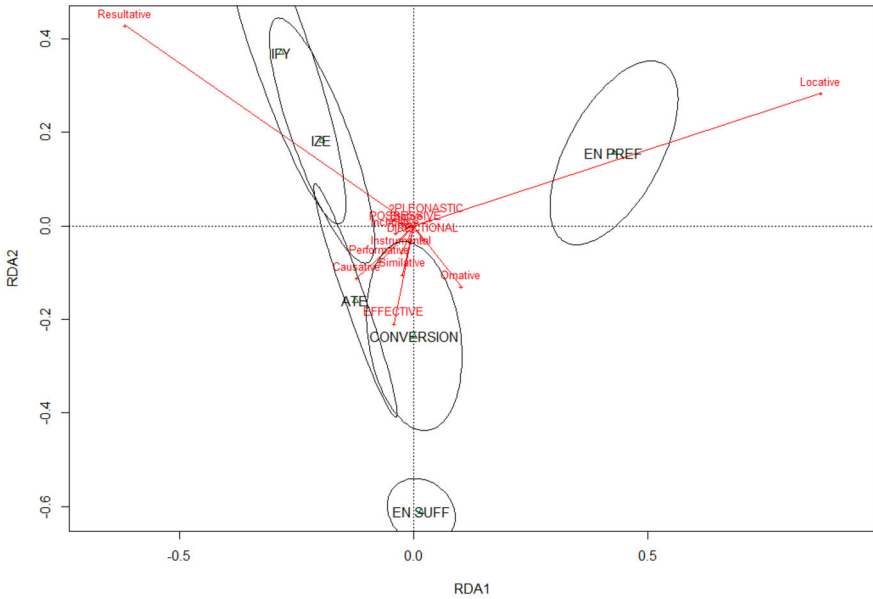
Figure 3: A representation of RDA for the English data sample.

    b.   *-ify* and *-ize/-ise* appear associated with RESULTATIVE, attested in 65.22 % of the *-ify* types and in 45.45 % of the *-ize/-ise* types.

    c.   *-en* suffixation appears moderately associated with EFFECTIVE (a category not listed in the standard set of semantic categories), attested in 18.18 % of the *-en* types.

    d.   Zero-derivation appears associated with ORNATIVE and with EFFECTIVE, attested in 24.49 % and 14.29 % of the zero types, respectively.

iv) The latter two statements suggest that an apparently neglected category such as EFFECTIVE is more relevant for zero-derivation than has been claimed in the literature.

For Spanish, RDA reveals a significant model such that the variables *Word-formation processes/Affix* and *Number of types* explain the variation of similarities between entries (F = 5.81, Df = 7–295, p-value = 0.001). Axes RD1 and RD2 explain the variation significantly (RDA 1: Variation explained = 5.5 %, F = 18.43, Df = 1–295, p-value = 0.001; RDA 2: Variation explained = 4.8 %, F = 16.07, Df = 1–295, p-value = 0.001). Only the variable *Word-formation processes/Affixes* explains similarities between entries significantly (F = 6.49, Df = 5–211, p-value = 0.001).
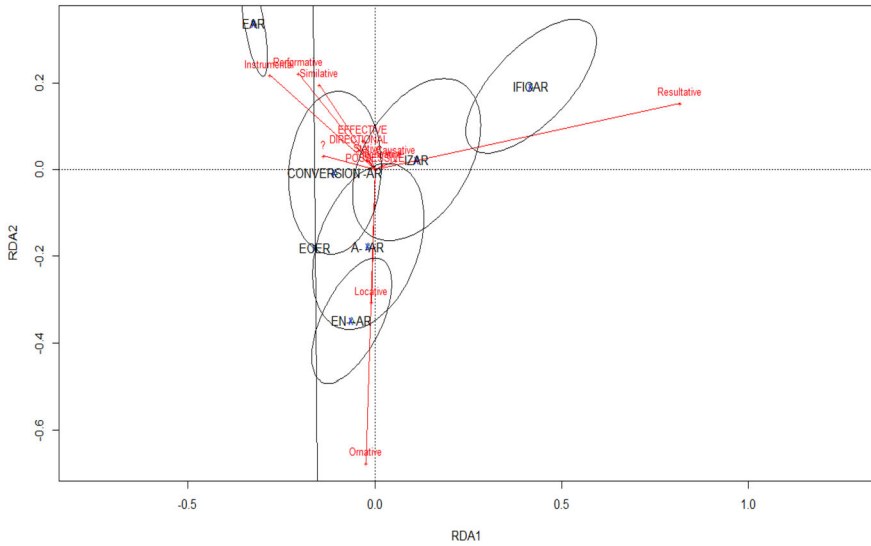
Figure 4: A representation of RDA for the Spanish data sample.

The variable *Number of types* does not yield a significant explanation and is discarded after forward selection. The final model after variable selection therefore relies only on the variable Processes, whereby the model remains significant (F = 6.49, Df = 5–211, p-value = 0.001).

RDA of the variable *Word-formation processes/Affix* accounts for 11.84 % of the variation by two axes, represented in Figure 4 (RDA 1: Variation explained = 5.5 %, F = 18.38, Df = 1–296, p-value = 0.001; RDA 2: Variation explained = 4.7 %, F = 15.75, Df = 1–296, p-value = 0.001).

Based on the above, the following claims can be made for Spanish:

i) All *Word-formation processes/Affixes* except -*(i)ficar* arrange themselves on a vertical continuum. Partial overlap occurs throughout, including outlier -*(i)ficar*, but not -*ear*, which is at one end of the continuum but does not overlap at all. Overlaps vary according to suffixes, e. g. -*izar*, *a-...-ar* and zero-derivation overlap considerably, whereas -*(i)ficar* and -*izar* overlap considerably less.

Suffixes cluster along a vertical band with -*ear* and *en-...-ar* at each end, and zero-derivation alongside the other affixes in between both ends, even if it lies closer to the end of circumfixation, i. e. closer to *a-...-ar* and *en-...-ar*.

ii) Circumfixation does not appear separate from the rest, even if the cases of circumfixation arrange themselves towards the lower end of the vertical band.

iii) Specific associations between affixes and categories appear, e. g.:

    a.   *a-…-ar* appears associated with the category ORNATIVE, attested in 62 % of the *a-…-ar* types. The same applies to *-ecer* (except that this affix is represented only by four low-frequency entries).

    b.   *en-…-ar* appears associated with the categories ORNATIVE and LOCATIVE, attested in 56 % and 26 % of the *en-…-ar* types, respectively.

    c.   *-izar* appears associated with ORNATIVE and RESULTATIVE, attested in 40 % and 30 % of the *-izar* types, respectively.

    d.   *-ear* appears associated with INSTRUMENTAL, PERFORMATIVE and SIMILATIVE, attested in 30 %, 24 % and 18 % of the *-ear* types.

    e.   Zero-derivation appears associated with the category ORNATIVE, attested in 42.86 % of the zero types. It shows moderate associations with PERFORMATIVE, INSTRUMENTAL, and RESULTATIVE, attested in 14.29 %, 10.20 % and 12.24 % of the zero types, respectively. It also appears moderately associated with a semantic category that does not fit any of the categories described in the literature for zero-derivation, whether in English or Spanish (and is therefore marked as OTHER[30]), attested in 6.12 % of the zero types. This suggests that the range of categories is in need of further development.

5 Discussion

Parallels between conversion and derivation are felt, to the extent that the two are perceived as described as similar operations in substance (e. g. Pellegrini and Montermini 2020). The results presented in Section 4 lend themselves to a number of comparisons between English and Spanish denominal verb formation by zero-derivation or by overt affixation in this regard:

i)   Affixes overlap variously for the expression of semantic categories. Zero-derivation arranges itself along the overlapping cline. In English, it arranges itself closer to suffixation and away from prefixation. Specifically, it overlaps substantially with *-ate* suffixation but not with the rest of suffixes. In Spanish, where denominal verb-forming prefixation is not possible, it arranges itself along a tighter gradient, and overlaps substantially with *-ecer* and *-izar* suffixation, and with *a-…-ar* circumfixation.

---

30 E. g. *estiar* 'become dry; wither, as in summer' (my translation), as in the concordance cited in note 29.

ii) Suffixes cluster along a vertical band, where zero-derivation takes a central position alongside other suffixes, both in English and in Spanish. In English, this band sets suffixation apart from prefixation, whereas in Spanish, where denominal verb-forming prefixation is not possible, suffixation is not separate from circumfixation.

iii) Zero-derivation shows specific associations with respect to several semantic categories. The associations are slightly different for English (ORNATIVE and, significantly, also moderately with EFFECTIVE) and for Spanish (ORNATIVE, and then PERFORMATIVE, INSTRUMENTAL, and RESULTATIVE, and a semantic category that does not fit any of the categories described in the literature, namely OTHER).

Corpus evidence for or against the separation between zero-derivation and overt affixation is rare. The paper's outlook for such evidence finds results that may be used for that purpose, especially as regards the contents of points i) and ii) above, but the fact remains that the separation can neither be discarded nor confirmed, because the data lend themselves to interpretation.

The results arise from a powerful database, quantitatively for the number of tokens researched, and qualitatively for the analysis of senses within types rather than of just types. Even so, research on the association between processes and the semantic categories that they express reveals a pattern in English where zero-derivation aligns itself with suffixation by contrast with prefixation, but not in Spanish, where it separates itself from circumfixation. At the same time, zero-derivation overlaps with suffixes for denominal verb formation (e. g. -*ate* in English, and -*ecer* and -*izar* in Spanish) and brings zero-derivation within a vertical band where it takes a central position both in English and Spanish, very much like any other denominal verb-forming suffix would.

If the separation between suffixation and prefixation shown for English is taken to be indicative of different processes (always as regards the representation of semantic categories), then the opposite, i. e. the occurrence of zero-derivation alongside suffixes, may be interpreted as a sign that it is closer to being part of the resources used by suffixation than to being a separate process. The picture is not entirely clear-cut in that -*en* suffixation is separate from the rest of suffixes, even if alongside the line of the vertical band. It is unclear how important the position of this suffix is, especially considering that, albeit significant, the corpus evidence for this type of formations is comparatively lower than in the rest of cases, and that it is reportedly no longer productive (cf. Table 2a and note 4, respectively). It is therefore also unclear how telling it is as regards what the most accurate interpretation of Figures 3 and 4 should be. Again, the theoretical position that views the process at issue as zero-derivation may argue that the view

in other terms than affixation, specifically as conversion, is called into question based on the central position that the data of zero-derivation take alongside other denominal verb-forming affixes, very much as supported by empirical research on other languages (cf. Melloni and Iordăchioaia 2020 on deverbal noun formation in Italian or Pellegrini and Montermini 2020 on Classical Greek). By contrast, the theoretical position that views this process as conversion may question the interpretation of the data in terms of a zero suffix, as suggested by Figures 3 and 4, based on the separation that the data of *-en* suffixation shows with respect to the rest of affixes.

No such similar analysis is possible in Spanish, where denominal verb formation by prefixation is not available. The alternative, comparison with circumfixation does not show the same results as the comparison between suffixation and prefixation in English: in the Spanish sample, zero-derivation is close to (actually overlaps in part with) circumfixation, but then so do other suffixes too (*-ecer*, *-izar*). In that respect, circumfixation appears to align itself with suffixation, and zero-derivation groups itself alongside them too.

6 Conclusion

Statistical analysis of a sense-sensitive (i. e. arranged by senses rather than by types) quantitatively and qualitatively powerful corpus-based database reveals similarities between various denominal verb-forming resources including zero-derivation, and their associated semantic categories in two morphologically different languages, English and Spanish. At the same time, it reveals differences too.

In English, zero-derivation aligns itself along a cline of suffixation. The cline lies away from prefixation. This lends itself to interpretation: if different processes as suffixation and prefixation occur separately, then the alignment of zero-derivation with suffixation may be viewed as evidence that it is more like the operation of a zero affix rather than as a separate process (conversion), at least as far as the association with certain semantic categories is concerned. In Spanish, where no word-class-changing prefixation is recorded, zero-derivation appears close to suffixation and circumfixation too. Again, this invites interpretation of the process as derivation by a zero affix rather than by a separate process (conversion), at least as regards their association with semantic categories.

At the same time, the cline of affixation reveals some degree of separation between affixes in the two languages. None of these separations is as marked as the separation of prefixation with respect to suffixation (and circumfixation in Span-

ish). The (degrees of) separation are therefore open to interpretation as essentially similar but also as different processes.

Aside from the issue of zero-derivation, both English and Spanish reveal the peculiarity of certain affixes and the overlap of zero-derivation with other suffixes. In the two languages, zero-derivation lies in a rather central position. Different profiles can be identified in each language, except that the difference may be only as a result of the different range of derivational possibilities that each language has.

Additional evidence from the opposite direction, i. e. deverbal nominalization (e. g. as in Iordăchioaia et al. 2020) and other related formations is necessary before firmer conclusions can be drawn. Until then, the evidence presented here for the two languages is engaging, but does not confirm the identification of this process with suffixation and, hence, does not prime one or the other concept. Interpretation as one or the other remains open, but the evidence brings to the fore new arguments in favor of the hypothesis of zero-derivation and against a separate process of conversion. Further evidence might come from the comparison of the degree of polysemanticity of every type sampled for each of the processes considered here and from the productivity of every type for each specific semantic category. As in this paper, the variation that can be noticed within affixation alone may not allow to expect that analysis to yield firm conclusions on the interpretation of this process as zero-derivation or as conversion, but it will likely result in better knowledge and, therefore, in better founded interpretations.

Table A1: English types.

en-	-ate	-en	-(i)fy	-ize/-ise	Ø
enact 747	motivate 1198	Hasten 480	exemplify 709	summarize 731	jail 960
empower 556	articulate 825	Heighten 472	quantify 511	minimize 523	hook 785
enlist 507	substantiate 313	Lengthen 381	unify 422	jeopardise 263	sanction 501
endanger 453	assassinate 288	Dishearten 28	typify 291	harmonise 160	bound 499
engulf 360	orchestrate 256	Unthreaten 13	personify 140	customise 109	outlaw 290
enshrine 349	degenerate 211	re-christen 7	codify 112	pressurise 92	weather 251
encircle 250	fabricate 188	Shapen 3	glorify 111	dramatise 73	blare 108
entrench 207	captivate 124	Unfrighten 3	objectify 43	synchronize 56	stable 71
embalm 152	urinate 94	Loaden 2	stratify 40	empathise 42	widow 43
enchant 137	gravitate 81	Barken 1	beautify 31	traumatise 33	gargle 27
enslave 101	copulate 64	Lashthicken 1	liquify 18	digitise 23	diagram 18
enmesh 72	instantiate 40		emulsify 16	productise 20	glass 13
emboss 61	officiate 37		scarify 12	magnetise 15	kickstart 10
entrap 57	fractionate 30		versify 10	fossilise 13	disciple 8
ensnare 52	levitate 24		gentrify 6	pasteurise 11	nightclub 7
envision 39	granulate 18		zombify 5	computerize 9	receipt 6
entreat 38	methylate 14		overqualify 4	atomize 8	jelly 5
entomb 30	alkylate 9		denitrify 3	tyrannise 7	bullock 4
enrapture 23	ulcerate 9		carbonify 1	lobotomise 6	lesion 4
emplace 20	oxygenate 6		desertify 1	decolonise 5	upwarp 4
embattle 19	flocculate 5		extensify 1	anthologise 4	decoy 3
encrypt 16	iodinate 4		porterfy 1	missionize 3	necklace 3
encash 14	hypothecate 3		re-notify 1	allegorise 3	spectacle 3
enshroud 12	hyphenate 3			parameterize 2	armlock 2
engrain 9	reformulate 3			productionise 2	croft 2

Table A1 (continued)

en-		-ate		-en	-(i)fy	-ize/-ise		∅	
embank	8	autoregulate	2			alphabetise	2	fowl	2
enamour	7	extubate	2			apostrophise	2	hill	2
encoffine	6	fabulate	2			diphthongize	2	monger	2
empanel	5	herniate	2			periodize	2	passage	2
enroll	5	machinate	2			rubberise	2	strop	2
encyst	4	rubricate	2			botanize	1	watermark	2
envenom	3	viticultivate	2			catheterize	1	ankle	1
enwrap	3	beneficiate	1			caramelize	1	blore	1
enjamb	2	centuriate	1			decolorize	1	clerk	1
encapsule	2	consultate	1			euphemise	1	copyright	1
enlace	2	dehalogenate	1			journalize	1	diecast	1
ensheathe	2	deprotonate	1			mathematize	1	environ	1
entrammel	2	fluoridate	1			narratize	1	galaxy	1
enflame	2	habilitate	1			opinionize	1	hair	1
enchase	2	ministrate	1			recategorize	1	insect	1
endebt	1	nodulate	1			restandardise	1	kitten	1
enfranchize	1	overcalculate	1			robotize	1	mouse	1
enfervour	1	performance-indicate	1			signalize	1	mustard	1
enstructure	1	personate	1			symptomatize	1	peacock	1
enwomb	1	ruinate	1					repin	1
emblaze	1	sensate	1					saint	1
embosome	1	sporulate	1					story	1
enwind	1	sulphonate	1					tumour	1
								womb	1
Aggregate	4344		3890	1391	2489		2240		3657

Table A2: Spanish types.

a-...-ar	en-...-ar	-ear	-ecer	-(i)ficar	-izar	Ø
447 asociar	489 enfocar	478 planear	5 desfavorecer	404 cuantificar	520 caracterizar	494 contactar
411 agrupar	365 enjuiciar	214 costear	2 frutecer	193 notificar	291 obstaculizar	384 premiar
374 afianzar	298 enlazar	106 golear	2 reflorecer	131 ejemplificar	139 colonizar	291 reinar
290 aunar	194 entonar	93 olfatear	1 fosforecer	130 testificar	87 teorizar	212 estacionar
231 aparcar	141 encarcelar	69 guerrear		112 dosificar	69 valorizar	160 torturar
192 alumbrar	89 enrollar	52 vadear		80 tipificar	55 ironizar	121 galopar
112 acondicionar	68 enarbolar	35 bracear		79 fructificar	43 profetizar	94 nevar
99 apresurar	54 encarrilar	26 espolear		75 escenificar	31 monitorizar	77 patentar
78 ahorcar	47 encasillar	23 burbujear		54 glorificar	26 caricaturizar	60 oxigenar
71 aquilatar	34 enlodar	20 puntear		43 mortificar	16 ruborizar	49 incordiar
50 arrostrar	30 embridar	17 cabildear		41 personificar	13 higienizar	39 fabular
40 abrochar	31 enraizar	14 trapear		40 tonificar	9 ideologizar	32 parcelar
33 aventajar	19 encolar	12 liderear		36 crucificar	10 culturizar	25 contrapesar
27 abrasar	17 enharinar	11 ventear		35 masificar	9 satelizar	21 rebuznar
22 amortajar	13 enlatar	9 lancear		28 mitificar	7 herborizar	16 cincelar
19 apestar	12 encorsetar	8 timonear		25 tecnificar	6 islamizar	13 alambrar
15 apuntillar	9 enlutar	7 prorratear		24 reunificar	5 coprotagonizar	11 juramentar
12 atornillar	8 ensalivar	6 limosnear		23 estratificar	5 victimizar	9 coreografiar
10 abotonar	7 envainar	5 contrapuntear		12 petrificar	4 carbonizar	8 regimentar
9 apalabrar	5 encajonar	5 turistear		11 nidificar	4 vandalizar	6 apellidar
6 amurallar	5 encizañar	4 guitarrear		10 versificar	3 gargarizar	5 aperturar
5 achicharrar	5 enfangar	4 rutear		8 vinificar	3 literaturizar	5 rebrillar
5 acojonar	4 emplatar	3 carrerear		7 momificar	3 novelizar	4 epilogar
5 apolillar	4 engolosinar	3 talonear		6 zonificar	2 avalizar	4 troquelar
4 abroncar	3 encabritar	2 carnavalear		5 fluidificar	2 dicotomizar	3 biografiar
4 amotinar	3 enfardar	2 flanear		5 gasificar	2 latinizar	3 ovillar

Table A2 (continued)

a-...-ar	n	en-...-ar	n	-ear	n	-ecer	n	-(i)ficar	n	-izar	n	Ø	n
ajardinar	3	enhornar	3	jaranear	2			pesificar	2	metropolizar	5	tintar	2
acongojar	3	entablillar	3	perifonear	2			ramificar	2	reanalizar	5	centellar	2
apeligrar	3	encaprichar	2	siluetear	2			nadificar	2	salmuerizar	4	escobillar	2
abocetar	2	encordar	2	versionear	2			vitrificar	2	arcangelizar	4	latigar	1
acalorar	2	enjaezar	2	artistear	1			calcificar	1	antipatizar	2	rebotear	1
acochinar	2	enmallar	2	bivaquear	1			estatificar	1	aluminizar	2	tertuliar	1
acristalar	2	envergar	2	carpinterear	1			gelificar	1	catectizar	2	agostar	1
aculturar	2	enarenar	2	chiripear	1			saponificar	1	deodorizar	2	azufrar	1
apiñar	2	encapillar	2	conventillear	1			clorificar	1	desenfatizar	1	bizcochar	1
acartonar	1	enchilar	1	cronistear	1			corporificar	1	dializar	1	cizallar	1
acatarrar	1	encorralar	1	fogonear	1			damnificar	1	destugurizar	1	deflacionar	1
achiquerar	1	encuestionar	1	gatillear	1			desertificar	1	esterizar	1	doctorar	1
acojinar	1	encuclillar	1	heraldear	1			emulsificar	1	guatemalizar	1	eficienciar	1
alanzar	1	enfajar	1	jopear	1			escarificar	1	genitalizar	1	estiar	1
alustrar	1	enfollonar	1	linternear	1			escorificar	1	iconizar	1	gloriar	1
amillarar	1	engrapar	1	pajear	1			historificar	1	indizar	1	honorar	1
aparrillar	1	engavillar	1	pesadillear	1			lotificar	1	madrigalizar	1	justiciar	1
apiadar	1	engripar	1	placear	1			nitrificar	1	mondrianizar	1	medicinar	1
aportillar	1	enlozar	1	ratear	1			periodificar	1	perspectivizar	1	paginar	1
aproar	1	enrabietar	1	retranquear	1			plantificar	1	realuminizar	1	polucionar	1
arrevistar	1	ensemillar	1	tabernear	1			preclasificar	1	repoetizar	1	recechar	1
asaetar	1	entrapajar	1	taponear	1			prosificar	1	semaforizar	1	resinar	1
atezar	1	entrojar	1	trialear	1			quilombificar	1	teologizar	1	tragediar	1
avinagrar	1	enyuntar	1	viñetear	1			sobrecodificar	1	zoologizar	1	vigiar	1
Aggregate 2607		1987		1256		10		1659		1391		2177	

References

Alemany Bolufer, José. 1920. *Tratado de la formación de palabras en la lengua castellana*. Madrid: V. Suárez.

Alvar Ezquerra, Manuel. [1994] 2015. *La formación de palabras en español*. 8th edn. Madrid: Arco.

Baayen, Harald R. 1994. Derivational productivity and text typology. *Journal of Quantitative Linguistics* 1. 16–34.

Baayen, Harald R. & Rochelle Lieber. 1991. Productivity and English derivation: A corpus-based study. *Linguistics* 29(5). 801–844.

Baeskow, Heike. 2020. Prominence in noun-to-verb conversion. In Lívia Körtvélyessy & Pavol Štekauer (eds.), *Complex words. Advances in morphology*, 82–99. Cambridge: Cambridge University Press.

Bagasheva, Alexandra. 2017. Comparative semantic concepts in affixation. In Juan Santana-Lario & Salvador Valera (eds.), *Competing patterns in English affixation*, 33–66. Bern: Peter Lang.

Bauer, Laurie & Rodney Huddleston. 2002. Lexical word-formation. In Rodney Huddleston & Geoffrey K. Pullum (eds.), *The Cambridge grammar of the English language*, 1621–1721. Cambridge: Cambridge University Press.

Bauer, Laurie & Salvador Valera. 2005. Introduction. In Laurie Bauer & Salvador Valera (eds.), *Approaches to conversion / zero-derivation*, 7–17. Münster: Waxmann.

Becker, Thomas. 1990. *Analogie und morphologische Theorie*. Munich: Wilhelm Fink.

Borcard, Daniel, François Gillt & Pierre Legendre. 2011. Unconstrained ordination. In Daniel Borcard, François Gillet & Pierre Legendre (eds.), *Numerical ecology with R*, 115–151. Dordrecht: Springer.

Carling, Gerd, Sandra Cronhamn, Love Eriksen, Robert Farren, Niklas Johansson & Joost van de Weijer. 2013. The cultural lexicon of Indo-European in Europe: Quantifying stability and change. In Guus Kroonen, James P. Mallory & Bernard Comrie (eds.), *Journal of Indo-European Studies* 65. 39–68.

Cetnarowska, Bożena. 1993. *The syntax, semantics and derivation of bare nominalisations in English*. Katowice: Uniwersytet Śląski.

Clark, Eve V. & Herbert H. Clark. 1979. When nouns surface as verbs. *Language* 55. 767–811.

Dahl, Eystein, & Antonio Fábregas. 2017. Zero morphemes. In *The Oxford research encyclopedia of linguistics*. Oxford: Oxford University Press.

Fábregas, Antonio. 2016. *Las nominalizaciones*. Madrid: Visor.

Gries, Steffan Th. 2003. *Multifactorial analysis in corpus linguistics: A study of particle placement*. London: Continuum.

Harley, Heidi. 2020. In defense of zeros. Paper presented at the workshop *Derivational Zero Affixes*, 53rd Annual Meeting of the *Societas Linguistica Europaea*, 26 August–1 September. https://osf.io/dqp3e/?pid=ufzq3.

Iordăchioaia, Gianina, Susanne Schweitzer, Yaryna Svyryda & María Camila Buitrago Cabrera. 2020. Deverbal zero-nominalization and verb classes: Insights from a database. *Zeitschrift für Wortbildung/Journal of Word Formation* 4(2). 120–142.

Karius, Ilse. 1985. *Die Ableitung der denominalen Verben mit Nullsuffigierung im Englischen*. Tübingen: Max Niemeyer.

Kastovsky, Dieter. 1994. Verbal derivation in English: A historical survey or much ado about

nothing. In Derek Britton (ed.), *English historical linguistics 1994. Papers from the 8th International Conference on English Historical Linguistics (8. ICEHL, Edinburgh, 19–23 September 1994)*, 93–117. Amsterdam: John Benjamins.

Lázaro Carreter, Fernando. 1980. Transformaciones nominales y diccionario. In Fernando Lázaro Carreter (ed.), *Estudios de lingüística*, 73–81. Barcelona: Crítica.

Leech, Geoffrey N. 1974. *Semantics*. Harmondsworth: Penguin.

Legendre, Pierre, Daniel Borcard, F. Guilaume Blanchet & Stéphane Dray. 2012. *PCNM: MEM spatial Eigenfunction and principal coordinate analyses*. R package version 2.1-2.

Marchand, Hans. 1969 [1963]. *The categories and types of present-day English word-formation. A synchronic-diachronic approach*. 2nd edn. München: Carl Beck.

Melloni, Chiara & Gianina Iordăchioaia. 2020. Zero suffix in English and Italian deverbal nouns. Paper presented at the workshop *Derivational Zero Affixes*, 53rd Annual Meeting of the *Societas Linguistica Europaea*, 26 August–1 September.

Oksanen, Jari, F. Guillaume Blanchet, Michael Friendly, Roeland Kindt, Pierre Legendre, Daniel McGlinn, Peter R. Minchin, et al. 2019. *Vegan: Community ecology package*. https://cran.r-project.org/package=vegan.

Pellegrini, Matteo & Fabio Montermini. 2020. The fuzzy boundary between affixation and conversion in Ancient Greek denominal verbs. Paper presented at the workshop *Derivational Zero Affixes*, 53rd Annual Meeting of the *Societas Linguistica Europaea*, 26 August–1 September.

Pena, Jesús. 1991. La palabra: Estructura y procesos morfológicos. *Verba* 18. 69–128.

Pena, Jesús. 1999. Partes de la morfología. Las unidades del análisis morfológico. In Ignacio Bosque Muñoz & Violeta Demonte Barreto (dirs.), *Gramática descriptiva de la lengua española, III*, 4305–4366. Madrid: Espasa.

Plag, Ingo. 1999. *Morphological productivity*. Cambridge: Cambridge University Press.

Plag, Ingo. 2018 [2003]. *Word-formation in English*. 2nd edn. Cambridge: Cambridge University Press.

Quirk, Randolph, Sydney Greenbaum, Geoffrey N. Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.

R Core Team (2019) *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rainer, Franz. 1993. *Spanische Wortbildungslehre*. Berlin: de Gruyter.

Rainer, Franz. 2016. Spanish. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: An international handbook of the languages of Europe, IV*, 2620–2640. Berlin: de Gruyter.

Sanders, Gerald. 1988. Zero derivation and the overt analogue criterion. In Michael Hammond & Michael Noonan (eds.), *Theoretical morphology: Approaches in modern linguistics*, 155–175. San Diego, CA: Academic Press.

Santiago Lacuesta, Ramón & Eugenio Bustos Gisbert. 1999. La derivación nominal. In *Gramática descriptiva de la lengua española, Vol. III*, directed by Ignacio Bosque Muñoz & Violeta Demonte Barreto, 4505–4594. Madrid: Espasa.

Savoy, Jacques. 2012. Authorship attribution: A comparative study of three text corpora and three languages. *Journal of Quantitative Linguistics* 19(2). 132–161.

Serrano-Dolader David. 1999. La derivación verbal y la parasíntesis. In *Gramática descriptiva de la lengua española, Vol. III*, directed by Ignacio Bosque Muñoz & Violeta Demonte Barreto, 4683–4755. Madrid: Espasa.

Štekauer, Pavol. 1996. *A theory of conversion in English*. Frankfurt am Main: Peter Lang.

Tournier, Jean. 2010 [1985]. *Introduction descriptive à la lexicogénétique de l'anglais contemporain*. 2nd edn. Paris & Genève: Champion-Slatkine.

Valera, Salvador. 2015. Conversion. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *An international handbook of the languages of Europe. Vol. I*, 322–340. Berlin: de Gruyter.

Valera, Salvador. 2021. Main morphological formal means (III): Approaches to conversion. In Antonio Fábregas, Víctor Acedo-Matellán, Grant Armstrong, María Cristina Cuervo & Isabel Pujol Payet (eds.), *The Routledge handbook of Spanish morphology*, 40–54. London: Routledge.

Varela Ortega, Soledad. 1990. *Fundamentos de morfología*. Madrid: Síntesis.

Venables, William N. & Brian D. Ripley. 2002. *Modern applied statistics with S*. Dordrecht: Springer.