Verginica Barbu Mititelu*, Svetlozara Leseva and Ivelina Stoyanova

Semantic analysis of verb – noun zero derivation in Princeton WordNet

https://doi.org/10.1515/zfs-2022-2017

Abstract: This study offers insights into the similarities and differences between the zero suffix and overt English suffixes involved in verb-to-noun and noun-to-verb derivation. It is based on morphosemantically related pairs of noun and verb senses released as a Princeton WordNet standoff file, which are annotated with a set of fourteen semantic relations further enriched with information about the affix(es) used in the derivation process. We compare the zero suffix and the overt suffixes with respect to their overall frequency in the dataset and their frequency as per semantic relation. We describe their semantics in terms of the relation between the base and the derived word senses, and of the semantic classes of words involved in affixal and zero derivation. We argue that the zero suffix is highly underspecified, occurring with all semantic relations, even though it manifests some preferences with respect to both the semantic relations expressed and the semantic classes of words it attaches to.

Keywords: English; overt suffixes; semantic groupings; semantic relations; zero derivation; zero suffix

1 Introduction

Zero derivation is the morphological process of creating new words by means of what is called a *zero suffix*,[1] which stands out among the other affixes as it lacks form. The process is also called conversion[2] by many language specialists. The existence of

1 We will use the singular form "zero suffix" throughout this paper as a generic term to refer to both the nominalizing and the verbalizing zero suffixes.

2 However, given the specific meaning that the term *conversion* has also acquired, referring to a process different from derivation (Valera 2014), we will not use it here, lest we should give rise to any confusion.

***Corresponding author: Verginica Barbu Mititelu**, Romanian Academy Research Institute for Artificial Intelligence, Bucharest, Romania, E-mail: vergi@racai.ro

Svetlozara Leseva and Ivelina Stoyanova, Institute for Bulgarian Language, Bulgarian Academy of Sciences, Sofia, Bulgaria, E-mail: zarka@dcl.bas.bg (S. Leseva), iva@dcl.bas.bg (I. Stoyanova)

this process, and, consequently, the legitimacy of the zero suffix are controversial (Borer 2013; Lieber 1992, 2004; Myers 1984; Pesetsky 1995; Plag 1999).

However, as shown by Bauer et al. (2013), the arguments against the existence of zero derivation are not completely convincing. Moreover, psychological experiments (Darby and Lahiri 2016; Fruchter et al. 2013; Pliatsikas et al. 2014) have brought arguments in favor of the existence of a derivational zero suffix. A further argument for its existence is that there are overt suffixes whose function is also shared by the zero suffix (Sanders 1988). Consider the pair $drive_V$ – $driver_N$ versus $ally_V$ – $ally_N$: the noun denoting the person performing the activity expressed by the verb is created by adding the suffix *-er* to the verb in the former case, and by zero derivation, i.e., by adding the zero suffix to the verb, in the latter. Various meanings associated with zero derivation were observed and described (Bauer et al. 2013; Cetnarowska 1993; Clark and Clark 1979; Lieber 2004; Plag 1999).

The zero suffix is characterized by high ambiguity. Semantically, there is a diversity of possible meanings the zero suffix can add to the base word: when forming zero-derived nouns, it may denote the person who does the activity expressed by the base word (see $ally_V$ – $ally_N$ above), the means by which this activity is done ($cover_{\text{V}}$ – $cover_{\text{N}}$), the action expressed by the verb ($drain_{\text{V}}$ – $drain_{\text{N}}$), etc. Morphologically, different parts of speech can be created (Schönefeld 2005): noun from verb – see examples above, verb from noun: $document_{\text{N}}$ – $document_{\text{V}}$, verb from adjective: $clean_{\text{Adj}}$ – $clean_{\text{V}}$, noun from adjective: $final_{\text{Adj}}$ – $final_{\text{N}}$. We focus here on the zero derivation involved in English verb-to-noun or noun-to-verb pairs of word senses as represented in our dataset. We adopt a semantic perspective, with the aim of answering the following research questions: what are the semantic regularities involved in zero derivation (i.e., meaning relations between the members of a pair) and what are the semantic conditions under which it occurs (i.e., the semantic classes the nouns and verbs belong to)?

Directionality in zero derivation is an ongoing research issue. While many researchers favor a directionality analysis, even if not all of them adopt a derivational approach (Arad 2005; Clark and Clark 1979; Kiparsky 1982; Lieber 1992, 2004), others support a non-derivational stance (Borer 2013; Farrell 2001). There are also those who adopt a bi-directional approach (Becker 1993; Tribout 2020). We adopt the view that the pairs related through zero derivation involve directionality. However, the direction of derivation is not encoded in the data we use. Moreover, it is not always straightforward to determine. This is why we cannot include directionality in our analysis but can only make judgments with respect to individual pairs of word senses and possibly in light of certain semantic patterns, such as the ones presented in Section 5.

We first present our dataset (Section 2) starting with an overview of the Princeton WordNet, its organizing principles and the morphosemantic relations

used to label the pairs in its standoff file.[3] We briefly outline the methodology, paying particular attention to the properties of the dataset, the general distribution of zero derivation and affixal derivation across the whole set of derivationally related pairs in the dataset which provides a background for the analysis of zero derivation, with no indication of the derivation direction. We also take into account the distribution of zero and affixal derivation per morphosemantic relation, which already offers some insight into the semantics of the two types of derivation. Section 3 contains an overview of the distribution of the five most frequent affixes for each morphosemantic relation with an eye on their semantic underspecification and their frequency for each possible sense (in terms of the verb and noun classes they connect and the relation expressed). Further semantic analysis of zero derivation and affixal derivation is presented in Section 4: for each morphosemantic relation we identify the most frequent combinations of noun senses and verb senses belonging to particular semantic classes (as expressed by their semantic primes, see Section 2.1) with zero derivation and/or with affixal derivation and remark on some tendencies that are observed in the data. In Section 5, a deeper look into the PWN structure helps us identify some semantic groupings (defined in the form of PWN subtrees) where pairs of morphosemantically related word senses tend to gather. Finally, we draw conclusions in Section 6 and envisage possible lines of research emerging from the observations of the data.

2 Dataset compilation and research methodology

Below we present the main resource employed in the study – Princeton WordNet, with a view to the data and features which we use for the dataset compilation. Further, we outline the methodology used in the analysis of the data.

2.1 Princeton WordNet

Princeton WordNet (PWN henceforth) is a lexical semantic network for English: its nodes contain word senses that form synonym sets (called synsets). A word may occur in the network several times, equal to its number of senses, be it polysemous or homograph. PWN contains nouns, verbs, adjectives and adverbs. Synsets are

3 See the morphosemantic database available at https://wordnet.princeton.edu/download/standoff-files.

interlinked by means of semantic relations.[4] In the case of nouns and verbs, these relations are mainly hierarchical: hypo-/hyperonymy, holo-/meronymy, troponymy. Each verb or noun synset is assigned a semantic prime that shows its membership to a relevant semantic hierarchy corresponding to a relatively distinct semantic field (Miller et al. 1990); hence a noun or a verb's prime relates the synset to a relevant semantic class. Although the distinction among primes is not straightforward and the meaning of a synset might share semantics with more than one semantic class, we take the data as provided, i.e., each synset is assigned one prime. Even though this approach might not reflect all the semantic distinctions of a synset's meaning, it is informative enough in terms of the semantics of (zero) derivation. Table 1 shows the semantic primes for nouns and verbs respectively.

Derived words are also linked to their base in PWN by means of a derivational relation. For example, the noun sense *net:2* with the gloss "a trap made of netting to catch fish or birds or insects" is derivationally related to the verb sense *net:4* with the gloss "catch with a net". Derivational relations were enriched with semantic information: a set of 14 semantic relations (Agent, Bodypart, By-means-of, Destination, Event, Instrument, Location, Material, Property, Result, State, Undergoer, Uses, Vehicle) was thus established (Fellbaum et al. 2009). Table 2 lists these relations together with a brief description of their semantics, an example and their frequency in the PWN standoff file (see the last column).

The definitions of morphosemantic relations is not made explicit in the data provided, but many relations are grounded in well established semantic roles (though not necessarily equated with them) or represent more general or more specific versions of such roles. The definitions in Table 2 were inferred on the basis of examining the semantic relation between the pairs of word senses related through them. Our own understanding of their semantics, first presented in Koeva et al. (2016), has since evolved, cf. the concise critical overview of the relations offered by Barbu Mititelu et al. (2021).

A concern raised with respect to the so-defined morphosemantic relations, which we are well aware of, is that the grounds for the selection of this particular inventory are not clear. As previously noted (Barbu Mititelu et al. 2021), not all relations seem to be equally justified, being too general or too specific. For instance, Vehicle may qualify as a kind of Instrument. Our reasons for considering the two relations independently is that Vehicle is defined quite specifically and the relevant noun senses fall into clear-cut semantic classes. We would rather avoid reassigning

4 In WordNet, semantic relations link synsets, i.e., meanings, while lexical relations link word forms; e.g., hyponymy, meronymy are semantic relations, while antonymy and synonymy are lexical relations (Miller et al. 1990).

Table 1: The 25 semantic primes for nouns and 15 semantic primes for verbs in PWN.

Noun primes		Verb primes
noun.act: acts or actions	**noun.phenomenon**: natural phenomena	**verb.body**: verbs of grooming, dressing and bodily care
noun.animal: animals		**verb.change**: verbs of size, temperature change, intensifying, etc.
noun.artifact: man-made objects	**noun.plant**: plants	**verb.cognition**: verbs of thinking, judging, analyzing, doubting
noun.attribute: attributes of people/objects		**verb.communication**: verbs of telling, asking, ordering, singing
noun.body: body parts	**noun.possession**: (transfer of) possession	**verb.competition**: verbs of fighting, athletic activities
noun.cognition: cognitive processes and contents		**verb.consumption**: verbs of eating and drinking
noun.communication: communicative processes and contents	**noun.process**: natural processes	**verb.contact**: verbs of touching, hitting, tying, digging
noun.event: natural events		**verb.creation**: verbs of sewing, baking, painting, performing
noun.feeling: feelings and emotions	**noun.quantity**: quantities and units of measure	**verb.emotion**: verbs of feeling
noun.food: foods and drinks		**verb.motion**: verbs of walking, flying, swimming
noun.group: groups of people or objects	**noun.relation**: relations b/n people/things/ideas	**verb.perception**: verbs of seeing, hearing, feeling
noun.location: spatial position	**noun.shape**: two and three dimensional shapes	**verb.possession**: verbs of buying, selling, owning
noun.motive: goals	**noun.state**: stable states of affairs	**verb.social**: verbs of political and social activities and events
noun.object: natural objects (not man-made)	**noun.substance**: substances	**verb.stative**: verbs of being, having, spatial relations
noun.person: people	**noun.time**: time and temporal relations	**verb.weather**: verbs of raining, snowing, thawing, thundering

this relation as Instrument so as to prevent losing the more specific information that is currently encoded. On the other hand, other relations, such as Uses, seem very abstract. Indeed, redefinition of some (instances of) relations is possible and meaningful and may be tackled in future research. However, our understanding is that the different granularity and specificity of morphosemantic relations is inherent and stems from the different granularity of prominent semantic components so reexamination of relations should be done with caution.

Table 2: The 14 morphosemantic relations used in the PWN standoff file with the number of available pairs.

Relation	Description	Example	Number
Agent	An entity that acts volitionally so as to bring about a result	*ruin – ruiner*	2,958
Body-part	A part of the body (e.g. of an Agent) involved in the situation	*extend – extensor*	69
By-means-of	Something that causes, facilitates, enables the occurrence of	*float – floater*	1,118
Destination	A recipient, an addressee or a goal	*patent – patentee*	28
Event	Something that happens at a given place and time	*beatify – beatification*	7,514
Instrument	An object (rarely abstract) acting under the control of an Agent	*instill – instillator*	782
Location	A concrete or an abstract place involved in the situation	*bifurcate – bifurcation*	273
Material	A substance or material used to obtain a certain effect or result	*sweeten – sweetener*	73
Property	An attribute or a quality	*magnetize – magnetization*	276
Result	The outcome of the situation described by the verb	*syllabify – syllable*	1,358
State	An abstract entity, such as a feeling, a cognitive state, etc.	*demoralize – demoralization*	514
Undergoer	An entity affected by the situation described by the verb	*invite – invitee*	847
Uses	A function an entity has or a purpose it serves	*attest – attestation*	726
Vehicle	An artifact serving as a means of transportation	*fight – fighter*	88

2.2 Dataset compilation

The dataset that we subjected to analysis consisted of a total of 16,624 verb – noun sense pairs obtained from the initial dataset included in the PWN standoff file (containing 17,741 pairs), after removing duplicate entries and errors in the relation annotation process and changing some relations for the purposes of consistency (the exact procedures, involving correcting duplicate morphosemantic relations and inconsistencies in the morphosemantic relations assigned are presented in detail by Koeva et al. (2016).

A verb – noun pair can occur several times in the PWN standoff file, each time at least one of the words having a different sense number. For example, the relation between the noun sense *net*:2 and the verb sense *net*:4 (presented above) is

Instrument.[5] Other senses of the two words are also considered derivationally related and an appropriate semantic relation is attached to them, as shown in Table 3.

Figure 1 shows the distribution of each relation in the dataset. We notice that Event accounts for a little less than half of the total number of pairs and Agent for almost a fifth of the pairs. The remaining relations are less frequent.

Overall, the proportion of zero derivation in our dataset is 46.5%, which shows that it is extremely productive in English (Plag 1999). The distribution of zero derivation and affixal derivation with respect to each of the 14 morphosemantic relations is presented in Figure 2, which confirms that zero derivation expresses a variety of meanings and that no specific one can be defined for it (Clark and Clark 1979; Plag 1999): all morphosemantic relations have cases of zero derivation.

The distribution of zero and overt affixal derivation per morphosemantic relation shows three types of situations:

1. a higher proportion of zero derivation – for Uses, Location, Undergoer, Property, and Result,
2. a higher proportion of affixal derivation – for Agent, Destination, Material, State, and
3. almost equal distribution – for Vehicle, By-means-of, Body-part, Event, and Instrument.

2.3 Methodology

We highlight that the data analyzed in this paper cannot be taken as representative of the English language; they are only representative with respect to PWN standoff

Table 3: Semantic relations attached to derivational relations.

Verb gloss	Verb	Semantic relation	Noun	Noun gloss
"catch with a net"	*net*:4	Instrument	*net*:2	"a trap made of netting to catch fish or birds or insects"
"yield as a net profit"	*net*:2	Result	*net*:3	"the excess of revenues over outlays in a given period of time (including depreciation and other non-cash expenses)"
"construct or form a web, as if by weaving"	*net*:3	Result	*net*:6	"an open fabric of string or rope or wire woven together at regular intervals"

5 The examples use the following notation: word followed by its sense number (for the respective part of speech in the case of homographs) from PWN version 3.1, which is available for querying at http://wordnetweb.princeton.edu/perl/webwn.

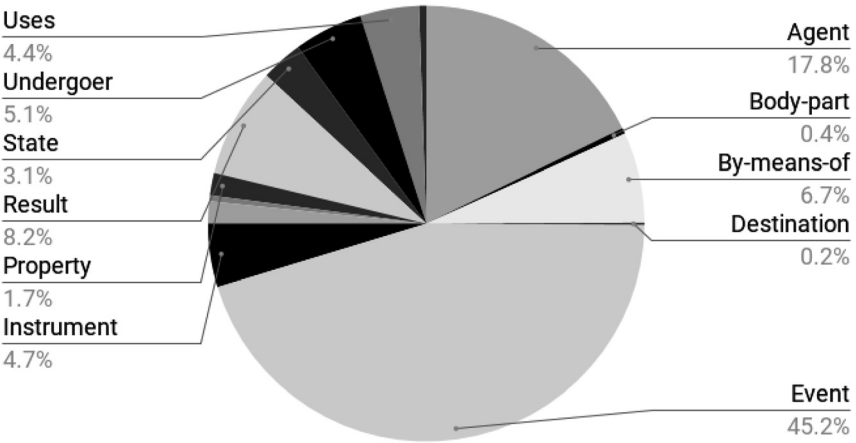

Figure 1: Distribution of semantic relations in the PWN standoff file.

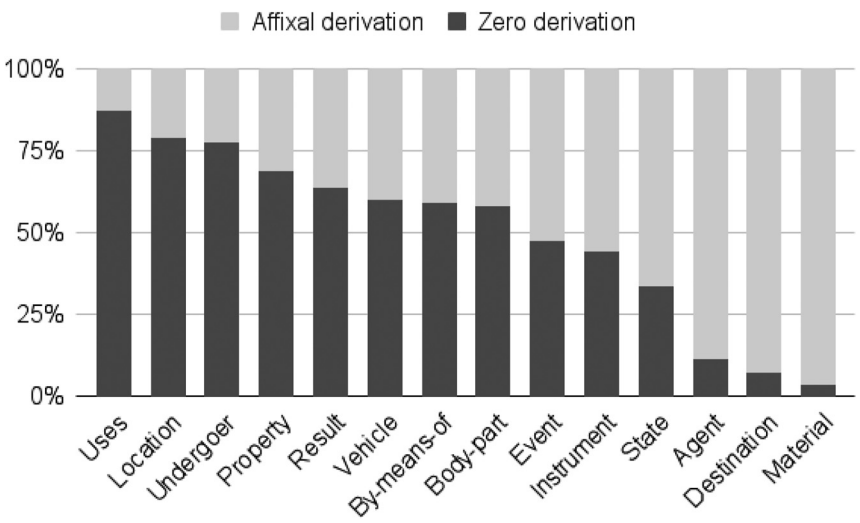

Figure 2: Distribution of derivation across the 14 morphosemantic relations.

file. To the best of our knowledge, documentation for the sampling of both derivational and morphosemantic relations has not been provided. We are not aware of the method adopted with respect to the selection of derivationally related word sense pairs that have been assigned morphosemantic relations. We can say, however, that the coverage is incomplete, as we have identified 4,520 verb – noun derivational pairs

that have not been assigned a morphosemantic relation (Barbu Mititelu et al. 2021), and it may also be the case that not all derivational relations existing between nouns and verbs have been identified and encoded as such in PWN. Even so, as PWN is a large-scale resource meant to represent the mental lexicon that has, in spite all the criticism, proved its usefulness in many computational linguistics tasks, we consider it a valuable dataset. Given the substantial number of pairs used in our analysis, we expect that the bias towards a particular type of derivation or part of the lexicon should be minimized.

We consider the distribution of zero derivation as a proportion of all the cases of derivationally related verb – noun sense pairs. As already discussed, we do not have information about the direction of derivation and thus, we do not factor it into the analysis, although we recognize the fact that it can be highly relevant. Moreover, resources containing directional information are limited (e.g., the Oxford English Dictionary covers only a portion of the derivationally related verb – noun pairs and does not always distinguish between different senses of the words).

The analysis in the following section is based on observations on the distribution of zero derivation across morphosemantic relations and in particular pays attention to the relations for which zero derivation dominates. A more in-depth analysis presented in Section 5 combines observations on the morphosemantic dimensions of derivation (including zero derivation) with analysis of the PWN tree structure and relations. For the identification of groupings we apply frequency analysis of different verb and noun subtrees and extract the corresponding subtrees with a concentration of morphosemantic links between synsets in them. We measure absolute (>5) and relative frequency (>10% of verb synsets) within a subtree entering a particular relation, and then we consider the larger groupings (of size >5) within subtrees in their noun counterparts. The automatic extraction is followed by manual analysis and selection. In order to facilitate automatic filtering and more precise identification, we need more detailed statistical analysis to discover statistically significant groupings. The present study serves as a preliminary attempt to provide some insights into the possible salient features and criteria in determining the groupings within the lexicon with respect to derivation and morphosemantics.

3 The distribution of zero and overt affixes

In this section we look at the semantics of the zero suffix, namely its meaning contribution to the sense of the derived word: this is reflected in the semantic relation that labels the pair of word senses that are morphologically analyzed as cases of zero derivation in our dataset.

For the purpose of this paper, for each pair in the standoff file we made explicit the affix(es) by means of which one word in the pair is obtained from the other. The identification of affixes was done automatically, using a manually compiled list of English prefixes and suffixes and a set of rules for attaching these affixes to the stem (e.g., transformation of final *y* preceded by a consonant into *i* before attaching the suffix, doubling the final consonant preceded by an unstressed vowel, etc.). The results of the automatic process were manually corrected to ensure an exact representation of the data. A first remark on this is that the nominal suffixes are much better represented than the verbal ones in out dataset. For each semantic relation a clear picture of the affixes involved and of their frequency emerges. Table 4 shows the most frequent suffixes for each relation.

The proportional contribution of each suffix to rendering the morphosemantic relations is represented in Figure 3. Each relation is dominated by one or two affixes at most and one of these is always the zero suffix (Ø). It is the most frequent suffix with the relations By-means-of, Body-part, Event, Location, Property, Result, Undergoer, Uses and Vehicle and the second most frequent with the relations Agent, Instrument, and State. For Destination it competes with the verb suffix *-ify* for the second place, while for Material it takes third place.

The zero suffix is the only one that can be found with all the relations, thus being the most underspecified among the suffixes in the dataset. As shown in Table 5, it has Event as default reading (45.41% of all its occurrences), but Result (with 11.34%), By-means-of (with 8.69%), Undergoer (with 8.52%) and Uses (with 8.40%) are also important readings. However, it is clearly not the only underspecified suffix, as other (noun) suffixes are also found with most of the relations, having various default and prevailing readings:

- *-ion* (including cases with *-(a)tion/-sion*) favors the Event relation (76.58% out of the 12 relations with which it is found),
- *-er* favors the Agent relation (76.34% out of the 12 relations[6] with which it is found),
- *-ment* is less underspecified (covering 10 relations) and has Event as its default reading (65.83%), while State (with 10.50%) and By-means-of (with 8.93%) are also available,
- *-ance*, *-al* and *-ee* are even less underspecified (covering only 8, 6 and 4 relations, respectively) with clear preference for particular relations – Event for *-ance* and *-al*, Undergoer and Destination for *-ee.*

As far as the verbal suffixes are concerned, Table 5 reveals that for both *-ify* and *-ise/-ize* the Result meaning prevails, although with the former this reading is more evident

6 The various readings of the suffix *-er* are also discussed by Fellbaum et al. (2009).

Table 4: The five most frequent suffixes for each morphosemantic relation. The zero suffix is rendered as Ø. Verb suffixes are labeled with a subscript *V*.

Relation	Ø%	1st	%	2nd	%	3rd	%	4th	%	5th	%	Rest
Agent	11.53	er	86.48	ion	0.68	ee	0.47	ise_V	0.24	ance	0.10	0.50
Body-part	57.97	er	39.13	ise_V	2.90	–	–	–		–		–
By-means-of	58.79	ion	16.64	er	14.58	ment	5.10	ance	1.25	ify_V	1.16	2.48
Destination	7.14	ee	85.72	ify_V	7.14	–		–		–		–
Event	47.22	ion	42.65	ment	5.59	ance	2.33	al	1.41	er	0.40	0.40
Instrument	44.08	er	52.94	ise_V	1.66	ion	0.77	ate_V	0.13	ment	0.13	0.29
Location	78.70	ion	8.79	er	5.13	ment	3.67	ise_V	2.56	–		1.15
Material	3.45	er	82.19	ise_V	9.59	ion	2.74	ate_V	1.37	–		0.66
Property	68.59	ion	20.25	ance	5.07	ment	4.51	ate	0.62	er	0.62	0.34
Result	63.45	ion	23.42	ise_V	3.98	ment	2.95	ify_V	2.95	al	0.81	2.44
State	33.40	ion	46.56	ment	13.04	ance	3.50	ise_V	1.56	al	1.56	0.38
Undergoer	77.73	ion	10.51	ee	4.25	ment	2.72	er	2.36	ance	1.18	1.25
Uses	87.43	ion	4.55	er	2.48	ise_V	1.38	ify_V	0.83	ment	0.69	2.64
Vehicle	59.78	er	40.22	–		–		–		–		–

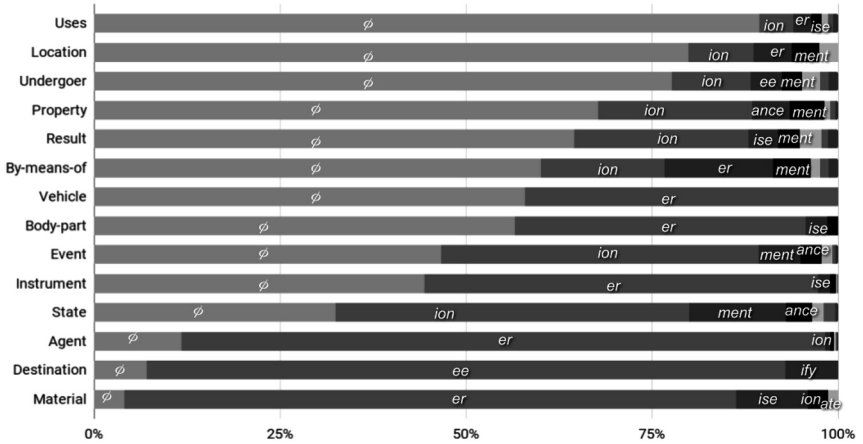

Figure 3: Distribution of suffixes within morphosemantic relations.

(55.55%) than with the latter (39.13%). Other notable meanings for *-ify* are By-means-of (18.05%) and Uses (8.33%). For *-ise/-ize* Event (13.77%), as well as Instrument (9.42%) and Uses (7.25%) are also important readings. For *-ate* the predominant reading is Result (26.67%), but there are also examples of Agent, By-means-of and Property (13.33% each).

4 Semantic primes: between zero derivation and affixal derivation

The semantic conditions under which zero derivation (Ø) and affixal derivation occur in terms of the semantic classes of the base and the derived word senses are presented below. In PWN, these semantic classes are defined by the semantic primes presented in Table 1 above. We thus use 'prime' as an abstract reference to the relevant semantic class when we study the regularities between the semantics of the verb and noun senses entering a particular morphosemantic relation.

Below we sum up the characteristics of each relation: these were identified by looking at the most frequent verb – noun prime pairs within the relation estimated in terms of the number of verb – noun pairs for each prime pair, and trying to further specify the relation's properties on the basis of the observed data. In addition, we looked at some clear-cut patterns that were identified in the data, and especially at those that were elucidating with respect to the semantics of the relation. We represent them by means of generalized definitions abstracted from the synsets' glosses, such as 'provide (with an) N'; these correspond to semantic patterns in the

Table 5: Distribution of relations for each of the most frequent suffixes. The zero suffix is rendered as Ø. Verb suffixes are labeled with the subscript *V*.

Relation/suffix	Ø	*-er*	*-ment*	*-ion*	*-ance*	*-al*	*-ee*	*-ise$_V$*	*-ify$_V$*	*-ate$_V$*
Agent	4.52	76.34	0.31	0.48	1.06	0.76	18.67	5.07	2.78	13.33
Body-part	0.50	0.80	–	0.02	–	–	–	1.45	–	–
By-means-of	8.69	4.86	8.93	4.45	4.93	–	1.33	5.80	18.05	13.33
Destination	0.03	–	–	–	–	–	32.00	–	2.78	–
Event	45.41	0.90	65.83	76.58	75.70	80.30	–	13.77	4.17	13.33
Instrument	4.49	12.35	0.16	0.14	–	–	–	9.42	–	6.67
Location	2.82	0.42	1.57	0.57	–	–	–	5.07	–	–
Material	0.04	1.79	–	0.05	–	–	–	5.07	–	6.67
Property	2.42	0.06	2.04	1.36	4.93	–	–	–	1.39	13.33
Result	11.34	0.24	6.27	7.60	2.46	8.33	–	39.13	55.55	26.67
State	2.16	–	10.50	5.83	6.34	6.06	–	5.80	2.78	–
Undergoer	8.52	0.60	3.61	2.13	3.52	3.79	48.00	2.17	4.17	–
Uses	8.40	0.54	0.78	0.79	1.06	0.76	–	7.25	8.33	6.67
Vehicle	0.66	1.10	–	–	–	–	–	–	–	–

sense intended by Marchand (1964) and to Plag's semantic representations (Plag 1999). Where possible in our analysis, we inspected all the relevant pairs, but where instances were too many (e.g. Event, with more than 8,000 verb – noun pairs), we looked at the overall tendencies.

Below, the relations are presented according to the prevalent type of derivation – zero derivation or affixation, with a third category reserved for those relations for which the distribution of these two types are rather similar.

4.1 Relations dominated by zero derivation

4.1.1 Uses

In many cases the noun corresponds to the Theme argument of the verb (an incorporated Theme encoded in the verb's root, Jackendoff 1990), as can be seen from the examples in Table 6.[7] Instances of this are predicates that refer to covering an object or a surface with an artifact (*carpet*:2) or a substance (*asphalt*:1). Other verbs have similar but more specific semantics: consider those whose definitions may be generalized as 'dress with an N' (*bonnet*:1, *jacket:2*), 'apply N' (*lipstick*:2), 'decorate

7 In the tables in this section, the data presented is not exhaustive, but represents only the most frequent prime pairs.

Table 6: Distribution of zero and affixal derivation for the Uses relation.

	Ø	Affixal	Examples
Distribution	87.43%	12.57%	
V.contact – N.artifact	101	6	*carpet*:2 – *carpet*:1, *wrap*:3 – *wrapper*:2
V.contact – N.substance	62	1	*asphalt*:1 – *asphalt*:1, *blot*:1 – *blotter*:1
V.possession – N.artifact	53	11	*armor*:1 – *armor*:1, *equip*:1 – *equipment*:1
V.communication – N.communication	39	8	*autograph*:1 – *autograph*:2; *attest*:1 – *attestation*:2
V.communication – N.cognition	4	4	*promise*:2 – *promise:*2; *exemplify*:2 – *example:1*

with N' (*garland*:1), among others. This explains the great frequency of V.contact verbs, which include, among other classes, verbs of putting and placing, on the one hand, and N.artifact and N.substance nouns (typical Themes), on the other. Other types of verbs may be defined as 'provide (with an) N' (*grate*:1, *evidence:2*), 'add N to' (*salt*:1), 'mark with an N' (*autograph*:1, *dot*:4), 'give (an) N', especially in a metaphorical sense: (*nickname*:1, *evidence*:3, *certificate*:1, *hurt*:2 'give trouble or pain to'), etc.

4.1.2 Location

The verbs denote the sense of 'move, place, maneuver, happen, etc. in/to/at N' or narrower classes that may be defined as 'live in N' *camp*:1 – *camp*:3, *bivouac*:1 – *bivouac*:1, etc. Some of the verbs may have a looser meaning varying from verb to verb, e.g. *cave*:2 'explore natural caves' – *cave*:1. The majority of noun senses are N.location, N.artifact and N.object denoting the location where the situation occurs or the goal that is to be reached (Table 7).

Table 7: Distribution of zero and affixal derivation for the Location relation.

	Ø	Affixal	Examples
Distribution	78.70%	21.30%	
V.contact – N.artifact	25	8	*cabin*:1 – *cabin*:1, *excavate*:4 – *excavation*:3
V.motion – N.location	23	1	*port*:2 – *port*:1, *culminate*:4 – *culmination*:2
V.contact – N.location	20	5	*border*:5 – *border*:1, *abut*:1 – *abutment*:1
V.change – N.location	6	6	*scour*:4 – *scour*:1; *locate*:4 – *location:1*,
V.change – N.artifact	4	4	*cloister*:3 – *cloister*:1; *compartmentalize*:1 – *compartment*:1
V.consumption – N.artifact	2	4	*mess*:1 – *mess*:5, *dine*:1 – *diner*:2

4.1.3 Undergoer

This relation shows great diversity with respect to both verb and noun semantic primes as the verb senses involving Undergoers form a broad and incoherent class, as do the noun senses that denote Undergoers (which broadly correspond to the Theme or Patient role or other affected entities in the theory of thematic roles). There are several combinations of primes occurring only with zero derivation, while N.person is the only N prime for which affixal derivation is more frequent than zero derivation (Table 8).

4.1.4 Property

The relation denotes the property, trait or quality associated with the verb sense and encoded in its semantics. In many cases this involves the change (including movement, change of location or position) of an entity so as to acquire a property. The most represented N prime is N.attribute in the most frequent data we have analyzed (Table 9).

4.1.5 Result

The relation is characterized by a similar diversity with respect to both N and V semantic primes as with Undergoers, for the same reasons – the broad scope of the activities and actions and the results yielded. Although the relation is clearly dominated by zero derivation, verbs of the type V.change show equal preference to

Table 8: Distribution of zero and affixal derivation for the Undergoer relation.

	Ø	Affixal	Examples
Distribution	77.73%	22.27%	
V.possession – N.possession	45	16	*store*:1 – *store*:2, *reimburse*:1 – *reimbursement*:1
V.communication – N.communication	39	31	*compliment*:1 – *compliment*:1, *communicate*:5 – *communication*:2
V.contact – N.artifact	33	4	*veneer*:1 – *veneer*:1, *wrap*:1 – *wrapper*:2
V.social – N.person	8	26	*outlaw*:1 – *outlaw*:1, *invite*:2 – *invitee*:1
V.competition – N.animal	16	0	*rabbit*:1 – *rabbit*:1
V.competition – N.artifact	10	0	*bomb*:1 – *bomb*:1
V.competition – N.food	5	0	*prawn*:1 – *prawn*:1
V.stative – N.artifact	6	0	*overhang*:2 – *overhang*:1
V.contact – N.animal	10	0	*snail*:1 – *snail*:2
V.contact – N.plant	8	0	*mushroom*:1 – *mushroom*:5
V.change – N.plant	11	0	*burr*:1 – *burr*:1

Table 9: Distribution of zero and affixal derivation for the Property relation.

	Ø	Affixal	Examples
Distribution	68.59%	31.41%	
V.change – N.attribute	32	25	*speed*:4 – *speed*:2, *appear*:2 – *appearance*:1
V.motion – N.attribute	19	2	*slant*:3 – *slant*:2, *concentrate*:5 – *concentration*:2
V.cognition – N.attribute	14	3	*distrust*:1 – *distrust*:2, *engross*:1 – *engrossment*:3
V.stative – N.attribute	11	9	*grace*:2 – *grace*:2, *abound*:1 – *abundance*:1
V.communication – N.attribute	10	7	*glamour*:1 – *glamour*:1, *deliberate*:2 – *deliberation*:5

both zero and affixal derivation, with some pairs of primes favoring one or the other type, as seen in the examples in Table 10.

4.2 Relations dominated by affixal derivation

4.2.1 Agent

Agent is the best represented relation for affixal derivation with sufficient number of attestations (Table 11). The vast majority of noun senses are of the type N.person with a small number of N.group, N.plant and N.animal, whereas the verb senses are quite diverse.

4.2.2 Material

This relation denotes a type of inanimate cause (Fellbaum et al. 2009) – substances that may bring about a certain effect. The prevalent prime pattern is V.change – N.substance, which accounts for the greater number of affixal pairs, but N.artifact

Table 10: Distribution of zero and affixal derivation for the Result relation.

	Ø	Affixal	Examples
Distribution	63.45%	36.55%	
V.creation – N.artifact	76	12	*corduroy*:1 – *corduroy*:2, *create*:2 – *creation*:2
V.contact – N.artifact	66	15	*bale*:1 – *bale*:1, *encrust*:2 – *encrustation*:3
V.communication – N.communication	40	25	*comment*:1 – *comment*:1, *declare*:6 – *declaration*:4
V.change – N.object	27	17	*crust*:1 – *crust*:2, *crystallize*:4 – *crystal*:3
V.change – N.shape	22	8	*curl*:1 – *curl*:1, *constrict*:2 – *constriction*:1
V.change – N.substance	20	44	*powder*:1 – *powder*:1, *calcify*:1 – *calcium*:1
V.change – N.state	6	25	*chafe*:1 – *chafe*:1, *refine*:1 – *refinement*:1
V.change – N.attribute	32	25	*color*:6 – *color*:1, *intensify*:4 – *intensity*:4

Table 11: Distribution of zero and affixal derivation for the Agent relation.

	Ø	Affixal	Examples
Distribution	11.53%	88.47%	
V.communication – N.person	49	475	*blabber*:1 – *blabber*:1, *accuse*:2 – *accuser*:1
V.contact – N.person	7	334	*butcher*:1 – *butcher*:3, *carve*:2 – *carver*:2
V.social – N.person	87	340	*chairman*:1 – *chairman*:1, *betray*:2 – *betrayer*:1
V.motion – N.person	23	272	*chauffeur*:1 – *chauffeur*:1, *commute*:2 – *commuter*:2
V.possession – N.person	22	226	*pirate*:1 – *pirate*:1, *auction*:1 – *auctioneer*:1

(synthetic substances or products) also qualify for this relation. As shown in Table 12, the second and third most represented pairs are much scarcer. In certain cases, with the meaning 'apply/cover with N' it may be considered as an instance of the relation Uses (and may be reassigned to it).

4.2.3 Destination

Destination is a relation between a verb sense and its recipient or addressee, on the one hand, or its destination (physical or abstract location), on the other. The former is better represented; all instances illustrate affixal derivation with the suffix *-ee* (see Table 4), and the nouns belong exclusively to N.person (Table 13). The zero derivation is instantiated by one example of V.cognition – N.group (*class:1 – class:7*) and one example of V.contact – N.artifact (*tee:1 – tee:3*).

Table 12: Distribution of zero and affixal derivation for the Material relation.

	Ø	Affixal	Examples
Distribution	3.45%	96.55%	
V.change – N.substance	1	62	*bronze*:1 – *bronze*:1, *coagulate*:1 – *coagulator*:1
V.body – N.artifact	1	8	*soap*:1 – *soap*:1, *carve*:2 – *carver*:2
V.contact – N.substance	0	5	*seal*:1 – *sealer*:1

Table 13: Distribution of zero and affixal derivation for the Destination relation.

	Ø	Affixal	Examples
Distribution	7.14%	92.86%	
V.possession – N.person	0	11	*grant*:7 – *grantee*:1
V.communication – N.person	0	8	*promise*:2 – *promisee*:1

4.2.4 State

State relates classes of verb senses associated with a resultative or another end state with noun senses denoting unspecified (N.state) or specified states (e.g. N.feeling, N.cognition) that come to exist or are involved in the situation described by the verb sense (Table 14).

4.3 Relations with a rather equal distribution of zero and affixal derivation

Several relations show relatively equal distribution between affixation and zero derivation. The most frequent combinations for them are presented below.

4.3.1 Event

For the Event relation zero and affixal derivation are in clear competition (Table 15). As one can notice, the pair V.communication – N.communication exhibits equal distribution of both derivational models (49% of cases for zero and 51% for affixal derivation), while most of the other frequent pairs favor one or the other. The overwhelming prevalence of the prime N.act (covering 52.75% of all cases) reflects its prototypical status for the relation as it expresses time- and place-bound acts and activities denoted by the respective verbs. Other very frequent primes are N.event (11.29%), N.process (5.84%) and N.state (3.58%), which also can refer to situations, as well as N.communication (13.40%) and N.cognition (5.52%), which represent more specific situations – related to communication or cognitive activities, events, and processes.

Table 14: Distribution of zero and affixal derivation for the State relation.

	Ø	Affixal	Examples
Distribution	33.40%	66.60%	
V.change – N.state	11	69	*silence*:1 – *silence*:1, *deforest*:1 – *deforestation*:1
V.emotion – N.feeling	33	45	*delight*:1 – *delight*:1, *captivate*:1 – *captivation*:2
V.emotion – N.state	20	30	*content*:2 – *content*:6, *afflict*:1 – *affliction*:1
V.social – N.state	14	31	*disgrace*:1 – *disgrace*:1, *imprison*:2 – *imprisonment*:2

Table 15: Distribution of zero and affixal derivation for the Event relation.

	Ø	Affixal	Examples
Distribution	47.22%	52.78%	
V.change – N.act	131	564	*lapse*:5 – *lapse*:2, *alter*:3 – *alteration*:2
V.communication – N.communication	324	337	*ban*:2 – *ban*:4, *examine*:4 – *examination*:1
V.social – N.act	201	386	*boycott*:1 – *boycott*:1, *assign*:1 – *assignment*:1
V.contact – N.act	335	220	*cuddle*:2 – *cuddle*:1, *affix*:2 – *affixation*:3
V.motion – N.act	391	111	*amble*:1 – *amble*:1, *arrive*:1 – *arrival*:2
V.communication – N.act	99	202	*dispute*:2 – *dispute*:2, *condemn*:2 – *condemnation*:2
V.change – N.process	34	253	*decay*:3 – *decay*:3, *adapt*:2 – *adaptation*:2
V.competition – N.act	119	47	*cricket*:1 – *cricket*:2, *invade*:3 – *invasion*:2

4.3.2 Instrument

For Instrument, the most frequent pair of primes (V.contact – N.artifact) is similarly active for zero and affixal derivation; apart from this pair and the affixation with the pattern V.change – N. artifact, the remaining combinations are much less represented (Table 16). Just like with Agent, the semantic range of the noun senses involved in the relation is very limited: mostly N.artifact (with occasional instances of N.communication), which points to the specificity of this relation.

4.3.3 Vehicle

N.artifact is the only noun prime that occurs in this relation. The most frequent verb prime is V.motion that makes up for most of the instances (79.35%). The other verb

Table 16: Distribution of zero and affixal derivation for the Instrument relation.

	Ø	Affixal	Examples
Distribution	44.08%	55.92%	
V.contact – N.artifact	224	168	*catapult*:1 – *catapult*:2, *dig*:1 – *digger*:2
V.change – N.artifact	23	110	*hamper*:2 – *hamper*:1, *freeze*:4 – *freezer*:1
V.motion – N.artifact	18	34	*centrifuge*:1 – *centrifuge*:1, *elevate*:2 – *elevator*:1
V.communication – N.artifact	13	27	*fax*:1 – *fax*:1, *buzz*: 1– *buzzer*:2
V.creation – N.artifact	26	20	*crayon*:1 – *crayon*:1, *cook*:2 – *cooker*:1
V.body – N.artifact	10	12	*comb*:3 – *comb*:3, *shave*:1 – *shaver*:3
V.competition – N.artifact	12	4	*harpoon*:1 – *harpoon*:1, *fend*:2 – *fender*:2

Table 17: Distribution of zero and affixal derivation for the Vehicle relation.

	Ø	Affixal	Examples
Distribution	59.78%	40.22%	
V.motion – N.artifact	53	20	*automobile*:1 – *automobile*:1, *cruise*:4 – *cruiser*:3
V.contact – N.artifact	2	9	*steamroller*:3 – *steamroller*:2, *intercept*:1 – *interceptor*:1
V.competition – N.artifact	0	6	*bomb*:1 – *bomber*:1

primes (V.contact, V. competition and V.creation) are few in number; they are presented in Table 17 as an illustration of the derivational patterns. The small number of primes and combinations reflects the very specific meaning of the relation, which, in more general accounts, would fall under Instrument; the meaning of the verbs entering the relation Vehicle is 'operate, travel, transport by means of N'.

4.3.4 Body-part

This relation is poorly represented in the dataset, which prevents us from making reliable judgments apart from noting the trend.

It denotes a part of the body that performs a certain action, e. g. *dilator*:1 or that someone uses to perform an action with, e.g. *finger*:1 and may be considered as a kind of Instrument that is a part of the Agent's body. Understandably, most noun senses involved are of the type N.body or N.animal, and the most frequent combination of primes is V.contact – N.body (Table 18). A couple of examples have different semantics, such as 'cover with N', e.g. *feather*:2, and may be reconsidered as instances of other relations.

4.3.5 By-means-of

The relation refers to a facilitating or mediating concrete or abstract entity (Means) which participates in the situation in a more general way than Instrument and

Table 18: Distribution of zero and affixal derivation for the Body-part relation.

	Ø	Affixal	Examples
Distribution	57.97%	42.03%	
V.contact – N.body	12	6	*finger*:1 – *finger*:1, *abduct*:2 – *abductor*:2
V.contact – N.animal	9	6	*paw*:1 – *paw*:1, *nip*:1 – *nipper*:2
V.motion – N.body	6	5	*elbow*:1 – *elbow*:1, *pronate*:1 – *pronator*:1
V.body – N.body	0	3	*tense*:3 – *tensor*:2

differs from Uses, which is directly involved (changes place, state, form, etc.) in the situation described by the verb. With verb senses such as *bridge*:3 "cross over on a bridge" or *railroad*:1 "transport by railroad", it links the verb to the artifact that serves as a means whereby the particular type of motion is carried out (Table 19). In other instances the relation may refer to an abstract relationship between a verb sense and something that facilitates the situation described. In addition, in rare cases the relation By-means-of may denote a causative meaning, e. g. *freeze*:6 'be very cold, below the freezing point' – *freeze*:2 'weather cold enough to cause freezing', *block*:6 'interrupt the normal function of by means of anesthesia' – *blocker*:2 'a class of drugs that inhibit (block) some biological process'. This calls for reconsideration and probably reassignment of some of the instances as well as, possibly, for a stricter reformulation of the relation.

5 Semantic groupings

In this section we touch upon a deeper level of analysis of verb – noun morphosemantics that emerges from the PWN data and builds upon prime combinations;[8] it may, however, be found implicitly in previous research as well. In addition to and within verb – noun prime pairs we identified more specific semantic patterns that often correspond to hierarchical substructures in the lexicon (organized roughly as

Table 19: Distribution of zero and affixal derivation for the By-means-of relation.

	Ø	Affixal	Examples
Distribution	58.79%	41.21%	
V.contact – N.artifact	115	27	*barricade*:3 – *barricade*, *impede*:2 – *impediment*:2
V.communication – N.communication	79	71	*alibi*:1 – *alibi*:1, *clarify*:1 – *clarification*:1
V.motion – N.artifact	29	13	*bridge*:3 – *bridge*:1, *canalize*:1 – *canal*:3
V.cognition – N.cognition	9	25	*estimate*:1 – *estimate*:3, *confirm*:1 – *confirmation*:2
V.communication – N.cognition	7	16	*lure*:1 – *lure*:2, *enlighten*:1 – *enlightenment*:1
V.change – N.artifact	20	16	*sandbag*:5 – *sandbag*:1, *fill*:1 – *filler*:4

8 As helpful as verb – noun pairs of primes are in defining and exploring morphosemantic relations, the level of abstraction of the primes is too general and most of the prime combinations are found across relations.

PWN hypernymic subtrees). Below we provide a glimpse into the potential of this kind of analysis although to paint a comprehensive picture would involve further analysis of the data.

Researchers have aligned derivational semantics (zero derivation in particular) with the semantics of the classes of verbs. Two such accounts are probably best known. Clark and Clark (1979), focusing on concrete nouns as bases for creating verbs by zero suffixation, describe five types of verbs that are further subdivided: locatum verbs (*blanket* the bed), location and duration verbs (*kennel* the dog, *summer* in Paris), agent and experiencer verbs (*butcher* the cow, *witness* an accident), goal and source verbs (*powder* the aspirin, *word* a sentence) and instrument verbs (*bicycle* into town). Different perspectives are assumed for semantically classifying nouns, too: for example, for agentive verbs the semantic classes of the parent nouns are occupations (*butcher* the cow), special roles (*referee* a game) and animals (*parrot* every word). Plag (1999) distinguishes among the following types of zero derived verbs, also suggesting a general semantic interpretation for each type: locative "put (in)to X" (*jail*), ornative "provide with X" (*staff*), causative "make (more) X" (*yellow*), resultative "make into X" (*bundle*), inchoative "become X" (*cool*), performative "perform X" (*counterattack*), similative "act like X" (*chauffeur*), instrumental "use X" (*hammer*), privative "remove X" (bark), stative "be X" (*hostess*).

The predefined 14 PWN morphosemantic relations determine similar or more specific semantic classes than the ones already described in the literature; some of the following non-exhaustive trends are observed. Clark and Clark's agentive and Plag's similative, which in essence is a kind of agentive, both correspond to the relation Agent; the same goes for the relation between instrumental verbs and their base nouns, on the one hand, and the Instrument morphosemantic relation, on the other. Locatum verbs correspond to the PWN subtree of *cover*:1 "provide with a covering" or *coat*:1 "cover the surface of" and possibly other subtrees; Clark and Clark's location/Plag's locative verbs are in line with the subtree of *put*:1. Ornative verbs may be aligned with the subtree of *supply*:1 "give something useful or necessary to". Locatum, location and ornative verbs, as described by the authors, are linked to their base nouns by means of the relation Uses in the account provided here.

As these examples show, given the hierarchical structure of nouns and verbs in PWN, the grouping of verbs and/or nouns with similar semantics correspond to hypernymic substructures in the PWN hierarchy. Thus, the morphosemantic relations may be viewed not just as connecting individual noun and verb senses, but as linking semantically coherent parts of the lexicon, i.e. subtrees. The PWN hierarchical structure allows us to observe more (often narrower) morphosemantically related formations in addition to the ones proposed by Clark and Clark (1979) and Plag (1999).

To illustrate this, we consider several such formations with the relation Undergoer. One is found between verbs of consumption and nouns denoting food and drink such as *drink*:1 – *drink*:1, *wine*:1 – *wine*:1, *claret*:1 – *claret*:2, *port*:7 – *port*:2, *tipple*:1 – *tipple*:2, *sup*:1 – *sup*:1, *taste*:3 – *taste*:5. These verb senses are all hyponyms of *consume*:2 "serve oneself to, or consume regularly" and all nouns denote the food or drink consumed, thus are hyponyms of *food*:1, so the subtree rooted in these two synsets (*consume*:2 and *food*:1) are a grouping in our understanding of the term. Within the same relation (Undergoer) we also find another such grouping with the hyponyms of *gather*:1 "assemble or get together" linked to hyponyms of *food*:1: *berry*:1 – *berry*:1, *blackberry*:1 – *blackberry*:1, *mushroom*:1 – *mushroom*:5, *oyster*:1 – *oyster*:2, *snail*:1 – *snail*:2 where the noun senses denote animals or plants serving as food that are being collected (the meaning encoded by the respective verb). A similar coherent grouping of verb – noun sense pairs is formed by hyponyms of *hunt*:1 "pursue for food or sport" and nouns denoting animals and/or fish, e.g. *whale*:1 – *whale*:2, *seal*:6 – *seal*:9, *fowl*:2 – *fowl*:1, *crab*:3 – *crab*:1, *shark*:2 – *shark*:1, etc. An interesting venue for further research would be a more detailed comparison with observations on such groupings provided in existing literature.

To analyze the regularities and interdependencies between derivational models (zero derivation in particular) and semantic relations between verb and noun senses that emerge from the data, we performed semantic grouping of verb – noun pairs for each morphosemantic relation. For a given relation we consider PWN verb (sub) trees and we investigate the verb senses that enter a morphosemantic relation R and their semantic primes (verb groupings). Secondly, in many verb trees for a relation R there are noticeable groups of verb senses which exhibit similar preferences for certain types of noun senses – in terms of their semantic prime and in terms of their place in the nouns' PWN tree (noun groupings).

Further, we illustrate how the study of morphosemantic relations between the members of subtrees reveals the morphological and semantic relatedness of hierarchically structured parts of the lexicon. In Figures 4 and 5 subtrees rooted in a given synset are represented as circles of different color and size, some of which are subsets of larger supersets. The numbers in the brackets in each set represent the number of pairs in the current set (the number on the left) out of the number of pairs in the closest superset that contains it (the number on the right).

In Figure 4 we show one semantic grouping for the relation Event with prime pairs displaying both zero and affixal derivation. We notice that different verb groupings tend to enter morphosemantic relations with particular noun groupings. For example, with regards to the Event relation the synsets in the tree of *communicate*:2 predominantly relate to nouns in the tree of *communication*:2, while verbs in the tree of *state*:1 relate to nouns in the subgroups of *speech act*:1 and *statement*:1.

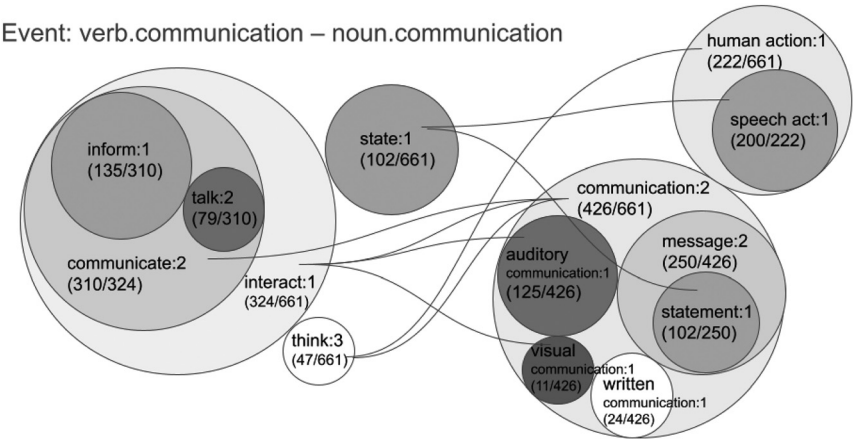

Figure 4: The interdependence between groupings for V.communication – N.communication with Event.

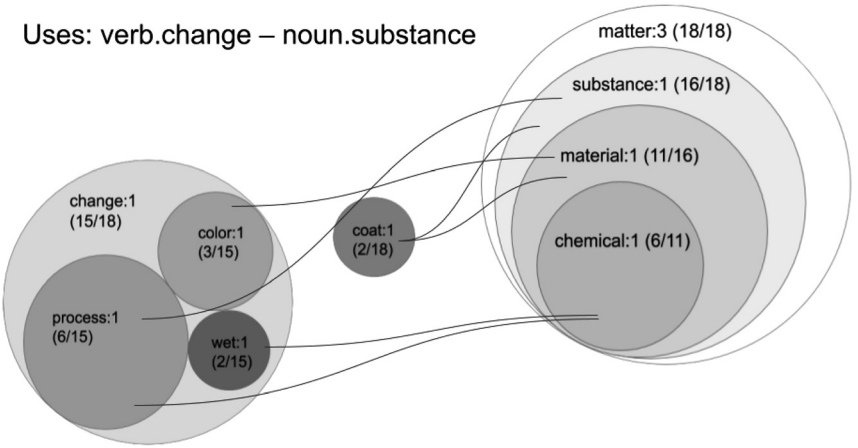

Figure 5: The interdependence between groupings for V.change – N.substance with Uses.

Similar trends are observed when we focus on groupings within zero-derivation pairs. In Figure 5 we show one semantic grouping for the relation Uses with pairs displaying zero derivation. For 18 pairs from our dataset, the noun subtree is rooted in the node *matter*:3 and the verbal one stems from the node *change*:1. We can notice that 6 verbs belong to the subsubtree rooted in *process*:1, while 16 nouns belong to the subsubtree rooted in *substance*:1. Out of these 16 nouns, 11 are hyponyms of *material*:1, and 6 of these 11 are grouped under the node *chemical*:1.

A further in-depth study into such groupings will provide a more profound look into the frequency and semantic patterns involved in zero derivation (with a recourse to affixation as well) and how they pattern across semantic classes. From a practical perspective these observations also help identify: (i) morphosemantic relations that are not typical for pairs of noun and verb subtrees, that is, another relation is prevalent between the members of these trees; hence, the untypical one may be a signal for a wrong automatic assignment: e.g., in the subtree of *cover*:1 there is a grouping of verbs related to nouns from the subtree of *food*:1 which in the majority of cases enter into the relation of Uses (*egg*:2 – *egg*:2, *flour*:1 – *flour*:1, etc.), with the exception of *butter*:1 where the relation assigned is Undergoer; thus Undergoer needs to be further validated or ruled out; (ii) relation gaps – missing morphosemantic relations between derivationally related verb – noun sense pairs identified in pairs of subtrees with groupings of morphosemantically related pairs: e.g., in the subtrees of the pair *communicate*:1 – *communicator*:1 where we observe 15 pairs representing the Agent relation, we also identify the derivationally related pair of synsets *supplicate*:1 – *supplicant*:2, *suppliant*:1, for which the relation Agent has not been encoded in PWN. These observations also facilitate the definition of more derivational models, involving a wider set of suffixes and taking into consideration phonetic transformations. These questions will be pursued in our future work.

6 Conclusions

The detailed analysis of the relationship between derivationally related word senses by taking into consideration their semantic primes in the PWN framework sheds light into the frequency of zero derivation versus affixation, into the frequency of zero derivation across relations, and of zero derivation with respect to pairs of semantic primes and smaller semantic groupings (PWN subtrees).

We have shown that the zero suffix is highly frequent in our dataset, highly underspecified with respect to the semantic relations it can express, and covers the whole spectrum of (14) relations defined over the dataset. It stands out among the affixes present in the data, as none of the overt ones, although sometimes highly underspecified too, may express all the relations. The zero suffix manifests preferences with respect to the relations expressed: its main meaning is Event, but Result, By-means-of, Undergoer and Uses are also important. At least in the data we worked with, it is also the main affix for creating new words for most of the relations (By-means-of, Body-part, Event, Location, Property, Result, Undergoer, Uses and Vehicle), it comes second in this respect with other relations (Agent, Instrument, State) and occupies the third position for the relation Material.

With respect to the semantic conditions under which zero derivation occurs, we have adopted a comparative approach between zero derivation and affixal derivation, and the most frequent pairs of semantic classes were presented for each morphosemantic relation and for each type of derivation. They show that although most of these combinations occur with both types of derivation, some of them are better represented for one type or the other. There are cases when, although zero derivation prevails with a certain relation, some pairs of semantic primes are more frequent with affixal derivation: see the pair V.social – N.person for Undergoer, or V.change – N.substance or V.change – N.state for Result: both relations are dominated by zero derivation, but these prime pairs are dominated by affixal derivation.

The pairs are extracted from a language resource (PWN) whose structure favors the analysis in terms of morphosemantically related groupings of words. They highlight semantic subtrees in which certain derivational tendencies can be noticed and, starting from them, undiscovered ones can be brought to light.

The results presented here are obtained by studying a large number of pairs of verbs and nouns. They offer more concrete arguments to already formulated linguistic characterizations of the phenomena under study, reveal certain tendencies and combinatorial preferences, as well as refine previous analyses. Although we cannot argue for the representativity of our PWN-based dataset for the entire English language, it is large enough to allow for generalizations about the semantics of zero and overt suffixes, which could be tested in future studies.

Further research should reconsider these results by adding another dimension to their interpretation, namely the direction of derivation.

Acknowledgments: We thank the anonymous reviewers and the editors for their feedback to previous versions of this paper, which helped us bring it to the current form.

References

Arad, Maya. 2005. *Roots and patterns: Hebrew morpho-syntax*. Dordrecht: Springer.

Barbu Mititelu, Verginica, Svetlozara Leseva & Ivelina Stoyanova. 2021. Semantic analysis of verb-noun derivation in Princeton WordNet. In *Proceedings of the Global WordNet conference 2021*, 108–117. Pretoria: Global Wordnet Association.

Bauer, Laurie, Rochelle Lieber & Ingo Plag. 2013. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.

Becker, Thomas. 1993. Back-formation, cross-formation, and 'bracketing paradoxes' in paradigmatic morphology. In *Yearbook of morphology*, 1–26. Dordrecht: Springer.

Borer, Hagit. 2013. *Taking form*. Oxford: Oxford University Press.

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(4). 767–811.
Darby, Jeannique & Aditi Lahiri. 2016. Covert morphological structure and the processing of zero-derived words. *The Mental Lexicon* 11(2). 186–215.
Farrell, Patrick. 2001. Functional shift as category underspecification. *English Language and Linguistics* 5. 109–130.
Fellbaum, Christiane, Anne Osherson & Peter E. Clark. 2009. Putting semantics into WordNet's 'morphosemantic' links. In Zygmunt Vetulani & Hans Uszkoreit (eds.), *Human language technology. Challenges of the information society. LTC 2007*, 350–358. Berlin, Heidelberg: Springer.
Fruchter, Joseph, Linnaea Stockall & Alec Marantz. 2013. MEG masked priming evidence for form-based decomposition of irregular verbs. *Frontiers in Human Neuroscience* 7. 1–16.
Jackendoff, Ray. 1990. *Semantic structures*. Cambridge, MA: MIT Press.
Kiparsky, Paul. 1982. From cyclic phonology to lexical phonology. In Harry van der Hulst & Norval Smith (eds.), *The structure of phonological representations*, 131–175. Dordrecht: Foris.
Koeva, Svetla, Svetlozara Leseva, Ivelina Stoyanova, Tsvetana Dimitrova & Maria Todorova. 2016. Automatic prediction of morphosemantic relations. In *Proceedings of the 8th Global WordNet conference (GWC)*, 169–177. Bucharest, Romania: Global Wordnet Association. Available at: https://aclanthology.org/2016.gwc-1.26.
Lieber, Rochelle. 1992. *Deconstructing morphology: Word formation in syntactic theory*. Chicago: University of Chicago Press.
Lieber, Rochelle. 2004. *Morphology and lexical semantics*. Cambridge: Cambridge University Press.
Marchand, Hans. 1964. A set of criteria for the establishing of derivational relationship between words unmarked by derivational morphemes. *Indogermanische Forschungen* 69. 10–19.
Miller, George A., Richard Beckwith, Christiane Fellbaum, Derek Gross & Katherine Miller. 1990. Introduction to WordNet: An online lexical database. *International Journal of Lexicography* 3(4). 235–244.
Myers, Scott. 1984. Zero-derivation and inflection. In Margret Speas & Richard Sproat (eds.), *MIT working papers in linguistics*, vol. 7, 53–69. Cambridge, MA: MIT Press.
Pesetsky, David. 1995. *Zero syntax: Experiencers and cascades*. Cambridge, MA: MIT Press.
Pliatsikas, Christos, Linda Wheeldon, Aditi Lahiri & Peter C. Hansen. 2014. Processing of zero-derived words in English: An fMRI investigation. *Neuropsychologia* 53. 47–53.
Plag, Ingo. 1999. *Morphological productivity: Structural constraints in English derivation*. Berlin/New York: 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: Academic Press.
Schönefeld, Doris. 2005. Zero derivation – functional change – metonymy. In *Approaches to conversion/zero derivation*, 125–155. Münster: Waxmann.
Tribout, Delphine. 2020. Nominalization, verbalization or both? Insights from the directionality of noun-verb conversion in French. *Journal of Word Formation* 4(2). 187–207.
Valera, Salvador. 2014. Conversion. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of derivational morphology*, 154–168. Oxford: Oxford University Press.