Verginica Barbu Mititelu, Gianina Iordăchioaia, Svetlozara Leseva & Ivelina Stoyanova

The meaning of zero nouns and zero verbs

Abstract: We carry out a large-scale study of noun-verb zero derivation pairs in English in order to identify possible semantic contrasts between the two derivational directions: V-to-N (zero nouns) and N-to-V (zero verbs). We compile a dataset of 4,879 N-V word sense pairs from the Princeton WordNet, which are annotated for noun and verb semantic classes and are assigned a morphosemantic relation. These sense pairs are labelled with a derivational direction from the Oxford English Dictionary. This makes it possible to investigate, on the one hand, the morphosemantic relations and, on the other hand, the noun and verb semantic classes that typically associate with each direction of zero derivation with the aim of offering a better understanding of the semantics involved in this morphological process.

Keywords: zero derivation, lexical semantics, derivational direction

1 Introduction

In this paper we are concerned with zero-derived noun-verb pairs in English as in (1), which we call *zero formations*, given the null marking of the categorial change undergone by the base.

(1) a. to climb > the climb (Zero nouns: V-to-N direction)
b. the bottle > to bottle (Zero verbs: N-to-V direction)

The choice of terminology depends on the theoretical view that linguists take on formations as in (1). While some consider these to instantiate a subtype of derivation which lacks morphological marking and call this process *zero derivation*, others consider them as representative of an independent word formation process which they call *conversion* to keep it separate from derivation (Valera 2014).

The goal of our study is to determine possible semantic differences between zero formations as in (1) depending on the direction in which they are built: i.e., from V to N as in (1-a) or from N to V as in (1-b). We work with a dataset of 4,879

Verginica Barbu Mititelu, Gianina Iordăchioaia, Svetlozara Leseva & Ivelina Stoyanova, VBM, Romanian Academy Research Institute for Artificial Intelligence, Bucharest; GI, Humboldt-Universität Berlin; SL & IS, Bulgarian Academy of Sciences, Sofia, email of corresponding author: vergi@racai.ro

https://doi.org/10.1515/9783111074917-004

N-V word sense pairs extracted from the Princeton WordNet standoff file containing pairs of morphosemantically related noun and verb senses (Fellbaum 1998; Fellbaum et al. 2009). For these pairs, we use the derivational direction provided by the Oxford English Dictionary (OED) (see §3.3) as a fixed predictor variable that will inform us on direction-specific semantic properties of zero formations. We investigate the semantic differences between zero nouns and zero verbs in terms of the semantic classes (or 'primes') of nouns (such as noun.artifact, noun.act, etc.) and verbs (verb.change, verb.motion, etc.) postulated by Miller et al. (1990) and the 14 morphosemantic relations (such as Event, Agent, Result) assigned to the N-V pairs in Fellbaum et al. (2009) (see §3.1). We undertake: (i) to seek whether such a large-scale study may reveal semantic properties typical of zero nouns and zero verbs and (ii) to check how our findings fare in relation to previous literature on the apparent differences between the two formations.

We address the following research questions:

1. Do the individual morphosemantic relations particularly associate with one derivational direction or the other? That is, do we find relations that primarily involve zero nouns or zero verbs?
2. Do the noun/verb semantic classes (or combinations thereof) particularly associate with either of the two directions in our dataset?
3. What do the apparent tendencies tell us about the semantics of zero nouns and zero verbs in view of previous theoretical studies?

A positive answer to the first two questions would allow an improvement of the current understanding of nouns and verbs as the output of zero derivation and, implicitly, of this morphological process so hotly debated in the linguistic literature. A qualitative (and quantitative) advantage of our approach is that it relies on a dataset compiled from large-scale resources available independently of our study, namely, the Princeton WordNet and the OED.

A close inspection of the morphosemantic relations and of the semantic classes ('primes') in our dataset allows us to identify some general semantic tendencies with respect to the derivational direction. On the one hand, we find that the morphosemantic relation Event favours zero nouns, relations such as Agent, Instrument, Location and others favour zero verbs, while relations such as State and Property are balanced between zero nouns and zero verbs. On the other hand, the majority of zero nouns instantiate the relation Event, while zero verbs are more evenly distributed between several such relations, indicating a more restricted pool of interpretations for zero nouns than for zero verbs. Moreover, within most relations, zero verbs exhibit more semantic classes than their base nouns, while zero nouns mostly show fewer or a similar number of semantic classes as their base verbs. This indicates that zero verbs expand the semantic range introduced

by the base, while zero nouns reduce it, pointing to a clear contrast between the two formations which also resonates with previous observations in the literature that we describe in Section §2.

We start with the theoretical background on zero nouns and zero verbs in Section §2. In Section §3 we present the lexical resources we used as well as the methodology for compiling our dataset of 4,879 zero-derived N-V word sense pairs and their analysis. Section §3.2 summarises our results, which we interpret in Section §5 in relation to our research questions and the previous literature. In Section §7 we conclude on our findings.

2 Theoretical background

In this section we briefly summarise the previous literature on zero nouns and zero verbs as a background against which we can evaluate our empirical findings. To our knowledge, the theoretical literature has not directly compared the two directions so far. However, an overview of the independent studies dedicated to each of the two formations, which we present in §2.1 and §2.2, suggests a contrast between the two derivational directions to the extent that zero nouns behave as expected of overt derivations by systematically restricting the semantics of the base, while zero verbs do not conform to this expectation. This contrast is confirmed by the computational study in Kisselew et al. (2016), which we summarise in §2.3.

2.1 Zero nouns in theoretical literature

Over the past few decades zero nouns have not received much attention in the literature, but they used to be more thoroughly investigated especially in the late 1960s (Kastovsky 1968; Marchand 1969; Irmer 1972). Marchand (1969) and Irmer (1972) analyse zero nouns as nominalisations by means of a zero suffix, and especially their semantics is described along the same patterns found with suffixed nominalisations: as action- and participant-denoting (see Plag et al. 2023) for a case study on the polysemy of *-ment* nominalisations).

Cetnarowska (1993) closely investigates the morphosyntactic and semantic behaviour of zero nouns and argues that they behave like deverbal nominalisations by means of suffixes such as *-ing*, *-ation*, *-ment*, *-al*, *-ance*. She argues that, like suffixed nominalisations, they are primarily action-denoting (e.g. *a jump*, *a pull*), while participant readings may appear more frequently than for suffixed nominals, but correspond to the same general patterns justified by meaning shifts:

result entities (e.g. *a cut*, *a rip*; cf. *a building*, *an establishment*), affected objects (e.g. *a find*, *a catch*; cf. *a borrowing*, *an inheritance*), and causers (e.g. *a bother*; cf. *a distraction*, *an embarrassment*, *a reminder*). For agents and instruments, zero nouns are in competition with the *-er* suffix (see *a guide* / *a guider*, *a sweep* / *a sweeper*), but typical action nouns may also form collective agents such as *administration*. For recent studies on affix competition in deverbal nominalisations see Plag et al. (2023) for English and Huyghe et al. (2023) for French.

Lieber (2016) also shows that zero nouns present semantic and morphosyntactic behaviour similar to that of suffixed nominals. Lieber (2016, 112) relies on her previous analysis of conversion as a conceptual process of lexical coinage, in which words are relisted with a different category (Lieber 1992). From this perspective, Lieber does not make any predictions about possible differences between zero nouns and zero verbs, even though she assumes directionality. However, she highlights important semantic parallelisms between zero nouns and suffixed nominalisations: see her parallel skeletons for zero and suffixed nominals (Lieber 2016, 111) and also her paper in this volume (Lieber 2023). Thus, Lieber's conclusions on zero nouns partially support their behaviour like overt derivations in relation to the base verbs.

This theoretical discussion indicates that zero nouns display substantial similarities to nominalisations derived with overt suffixes, as their semantics relies on systematic patterns defined in relation to their base verbs similar to those of suffixed nominals. Borer (2013, ch. 7) partially argues against such a similarity, although she does not address the polysemy of zero and suffixed nominals as Cetnarowska (1993) and Lieber (2016) do. Borer discusses argument structure, which she claims zero nouns fail to inherit from their base verbs, in contrast to suffixed nominals. However, her arguments have been seriously challenged by Lieber (2016) and Iordăchioaia (2020).

2.2 Zero verbs in theoretical literature

Zero verbs have been more at the centre of theoretical research. Clark & Clark (1979) insightfully show how spontaneous and productive denominal zero verb formation is and how these verbs may acquire a whole array of context-dependent and pragmatically-motivated meanings, which could not be predicted from the meaning of the base noun alone (see *He wristed the ball over the net*, stated by a tennis commentator). This finding cannot be accounted for by a directional derivational approach, which makes the centre of Lieber's (1981; 1992) proposal that zero verbs cannot be the output of a zero derivation process but must be a case of lexical coinage and relisting. Kiparsky (1982) and Hale & Keyser (2002) attempt to

identify regular patterns in the behaviour of zero verbs depending on the meanings of their base nouns. Kiparsky argues that some zero verbs incorporate the meanings of their base nouns as arguments in their event structure, while others do not. He distinguishes between instrumental verbs like *to hammer* and *to tape* as to whether they allow an adjunct PP that introduces an instrument different from that expressed by their base nouns, as in (2):

(2) a. Lola hammered the metal (with her shoe).
b. Lola taped pictures to the wall (*with pushpins).

From (2) Kiparsky concludes that *hammer*-type verbs are not derived from the base nouns, and their meanings are not tied to these, given that other instruments are possible in (2-a). *Tape* verbs, however, must be derived from their base nouns, since no other instruments are possible in (2-b) (see also Arad 2005).

Harley & Haugen (2007) challenge Kiparsky's contrast and argue that even *tape* verbs allow different instrument PPs to the extent that they involve the same manner of action, as illustrated in (3). The difference between band-aids and pushpins is that they involve different manners of action. Harley and Haugen conclude that not even verbs like *tape* encode the instrument of their base nouns in their meaning; the base nouns loosely specify the manner of action the zero verbs denote.

(3) Lola taped the poster to the wall with band-aids / mailing-labels.

Rimell (2012) provides extensive support for this conclusion on the basis of corpus data. In simple terms, she argues that denominal zero verbs cannot be viewed as regularly derived from their base nouns, since the meaning of the noun is not encoded in that of the verb; the base noun is interpreted as a predicate of events and not as an argument of the verb (cf. Hale & Keyser 2002).

Additional observations compatible with this observation are provided by Plag (1999) and Lieber (2004). While rejecting a treatment of zero verbs as the output of a zero derivation process, Plag argues that the semantics of a purported zero suffix for such verbs is too polysemous to instantiate one unitary suffix. For instance, he finds that of the 488 recent zero verbs in his dataset only 79 express meanings that are also associated with overt verbalising suffixes such as *-ise*, *-ify* or *-ate*. Plag concludes that a zero verbalising suffix would have no corresponding overt analogue (Sanders 1988) and is thus untenable.

From this theoretical overview we conclude that the meaning patterns that zero nouns acquire in relation to their base verbs represent a well-defined set similar to those of overtly derived nominalisations. By contrast, the meaning patterns that zero verbs receive are richer than those of overt verbalising derivations and

do not establish a systematic semantic relationship with the base such as that between a predicate and its argument. The meaning of the base only loosely contributes to some pragmatically and contextually determined manner specification on the event of the output zero verb. This indicates that the meaning of zero nouns is more restricted in relation to that of the base than the meaning of zero verbs is, a conclusion that is also reached by the computational study in Kisselew et al. (2016), which we summarise below.

2.3 Computational insights on zero nouns vs. zero verbs

For a manually created dataset of N-V pairs, Kisselew et al. (2016) use historical precedence as a proxy for the derivational direction and test the reliability of frequency and semantic specificity in this respect. To the extent that more directionality tests converge, this enforces the direction and shows that such N-V pairs are well-behaved. Semantic specificity refers to the observation that the output of overt derivation is semantically more complex and more specific than its base (see Koontz-Garboden's 2007 Monotonicity Hypothesis for derivation). If zero derivation is similar to overt derivation, zero verbs and zero nouns are expected to be semantically more specific than their base nouns and verbs.

Kisselew et al. apply measures of information content to distributional representations to estimate semantic specificity. For overt derivation, these measures yield 90% accuracy (Padó et al. 2015), confirming the validity of semantic specificity in this morphological process. While Kisselew et al.'s best model for zero derivation combines semantic specificity with frequency, semantic specificity alone yields almost as good results for zero nouns and drops to chance level for zero verbs. This means that semantic specificity successfully predicts zero nouns (like in the case of overt derivations) but fails with zero verbs. Kisselew et al. take this result to indicate that zero nouns are semantically more specific than their bases and consistent with a derivational approach, while zero verbs are semantically more irregular and compatible with a non-derivational approach (see overview on approaches to zero derivation in Darby (2015, §1.4).

While the theoretical studies in §2.1 and §2.2 have not directly addressed semantic specificity as used in Koontz-Garboden (2007) and Kisselew et al. (2016), their independent insights on the semantics of zero nouns and zero verbs converge with those in Kisselew et al. (2016) and support the contrast the latter observe. In §4.3 we will have a look at the semantic classes that nouns and verbs show in our dataset when they represent the base or the output of zero derivation and we will see that our PWN data provide some support for this contrast between zero nouns and zero verbs, as well.

3 Resources and methodology

Below we briefly describe the resources we used in our analysis. The Princeton WordNet provides the initial set of noun-verb sense pairs annotated with morphosemantic relations and the semantic class each word sense belongs to. From the OED we draw information about the direction of zero derivation. To serve our objectives, the two resources were aligned with respect to the data under focus.

3.1 Princeton WordNet

Princeton WordNet (PWN) was conceived as a network of English word senses: its nodes are represented by synonym sets (called synsets) of words of the same part of speech, be they nouns, verbs, adjectives or adverbs. A word may occur in the network several times, equal to its number of senses, be it polysemous or homograph. Synsets are interlinked by means of a number of semantic relations: within the classes of nouns and verbs these relations are mainly hierarchical: hypo-/hyperonymy, holo-/meronymy, troponymy. Each verb or noun synset is assigned a semantic prime (Miller et al. 1990), which shows its membership to a relevant semantic category; hence a prime assigns a noun or a verb to a relevant semantic class. Table 1 shows in boldface the semantic primes for nouns and verbs. Each semantic class is organised into one or more trees in PWN. 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 they are provided, i.e., each synset is assigned one prime.[1] This approach might not reflect all the semantic distinctions of a synset's meaning, but it is informative enough in terms of the semantics of (zero) derivation. Further considerations on the representativeness of the data are given in §3.4.

Besides the conceptual relations between synsets (e.g. hypo-/hyperonymy), there are also lexical relations between words from different synsets in PWN. Derivational relations are instances of such lexical relations. For example, the noun *cost:1* with the gloss "the total spent for goods or services including money and time and labor" is derivationally related to the verb *cost:1* with the gloss "be priced at". Note, however, that, as implemented in the PWN, these relations do not show the direction of the derivation, but only the lexical items involved. In our

1 The PWN lexicographic files are downloadable from https://wordnet.princeton.edu/download.

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

endeavour, labelling the direction has necessitated the additional employment of the OED (see §3.2).

Fellbaum et al. (2009) enriched these derivational relations with semantic information leading to a set of 14 morphosemantic relations: Agent, Body-part, By-means-of, Destination, Event, Instrument, Location, Material, Property, Result, State, Undergoer, Uses, Vehicle. These morphosemantic relations were designed following previous lexical semantic literature (see the Cases proposed in Fillmore 1968, the frame elements of FrameNet in Ruppenhofer et al. 2002 and the semantic roles in Gildea & Jurafsky 2000) and we used them to annotate a set of 17,739 noun-verb pairs (irrespective of the direction of the derivational process) made available

as a standoff file. Table 2 lists these relations together with a brief description of their semantics and examples from PWN.

Tab. 2: The 14 morphosemantic relations used in the PWN standoff file.

Relation	Description	Example
Agent	an entity that acts volitionally so as to bring about a result	*ruin – ruiner*, *bully*$_N$ – *bully*$_V$
Body-part	a part of the body (e.g. of an Agent) involved in the situation	*extend – extensor*, *finger*$_N$ – *finger*$_V$
By-means-of	something that causes, facilitates, enables the occurrence of	*float – floater*, *smell*$_V$ – *smell*$_N$
Destination	a recipient, an addressee or a goal	*patent – patentee*, *tee*$_N$ – *tee*$_V$
Event	something that happens at a given place and time	*beatify – beatification*, *chat*$_V$ – *chat*$_N$
Instrument	an object (rarely abstract) acting under the control of an Agent	*instill – instillator*, *microwave*$_N$ – *microwave*$_V$
Location	a concrete or an abstract place involved in the situation	*bifurcate – bifurcation*, *chamber*$_N$ – *chamber*$_V$
Material	a substance or material used to obtain a certain effect or result	*sweeten – sweetener*, *plaster*$_N$ – *plaster*$_V$
Property	an attribute or a quality	*magnetise – magnetisation*, *overlap*$_V$ – *overlap*$_N$
Result	the outcome of the situation described by the verb	*syllabify – syllable*, *ash*$_N$ – *ash*$_V$
State	an abstract entity, such as a feeling, a cognitive state, etc.	*demoralise – demoralisation*, *pother*$_N$ – *pother*$_V$
Undergoer	an entity affected by the situation described by the verb	*invite – invitee*, *harvest*$_N$ – *harvest*$_V$
Uses	a function an entity has or a purpose it serves	*attest – attestation*, *brine*$_N$ – *brine*$_V$
Vehicle	an artifact serving as a means of transportation	*cruise – cruiser*, *sled*$_N$ – *sled*$_V$

In the PWN standoff file, a pair of two words can occur several times, each time at least one of the words having a different sense number. For example, the relation between the noun *cost*:1 and the verb *cost*:1 is Event.[2] Other senses of the two

2 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.

words are also considered derivationally related, and an appropriate morphosemantic relation is attached to them, as shown in Table 3.

Tab. 3: morphosemantic relations attached to derivational relations.

Verb gloss	Verb	morphosemantic relation	Noun	Noun gloss
"be priced at"	*cost:1*	Event	*cost:1*	"the total spent for goods or services including money and time and labor"
"be priced at"	*cost:1*	Property	*cost:2*	"the property of having material worth (often indicated by the amount of money something would bring if sold)"
"require to lose, suffer, or sacrifice"	*cost:3*	Result	*cost:3*	"value measured by what must be given or done or undergone to obtain something"

As many morphosemantic relations are grounded in well-established generalisations of semantic roles, they are liable to the same line of criticism. One of the concerns raised, which we are well aware of, is that the grounds for the selection of this inventory are not clear. In particular, some relations may cover distinct meanings, and not all relations seem to be equally justified or of the same level of granularity. An analysis of the morphosemantic relations with proposed definitions and semantic restrictions (couched in terms of noun and verb primes) is presented in Koeva et al. (2016). A concise critical overview in light of further observations on the relations is dealt with in Mititelu et al. (2021). Our understanding is that the differences among morphosemantic relations are inherent to and stem from the varying granularity of prominent semantic components even if some of them may be redundant for our task. Therefore, we take the data as given and try to make sense of the generalisations underlying the distinctions made in the resource, where relevant.

3.2 Oxford English Dictionary

The OED contains 600,000 words from over 1,000 years of history of the English language. For each word in our dataset, we used the OED API to obtain the list of lemmas (separate OED entries) with the part of speech and the list of senses with definitions.

In addition, we have applied two special resources derived from OED:[3]

- a list of 2,830 pairs of verb and noun OED lexical entries labeled as V-to-N zero derivation;
- a list of 5,921 pairs of verb and noun OED lexical entries labeled as N-to-V zero derivation.

These were processed into 2,660 **unique** word pairs V-to-N, 5,521 **unique** word pairs N-to-V and 347 **unique** word pairs that show both directions of derivation in some of their senses.

3.3 Dataset compilation

Our data initially comprised the overall dataset of verb – noun pairs from the PWN that are labeled with a morphosemantic relation (see §3.1). For each pair we assigned the direction of derivation if known from the OED (V-to-N, N-to-V or both). We discarded pairs if (i) there are homonym entries in OED for some of which a direction is not assigned (so we do not know whether our WordNet senses correspond to any of the relevant OED senses and whether the derivational direction stands), or if (ii) two different directions are assigned to different senses of the verb and noun, so no clear direction can be established between the two.

We also make sure that the WordNet gloss is similar to one or more of the OED senses in order to confirm the reliability of the directionality information transferred from the OED onto the WordNet entries. The verification was initially performed by measuring text similarity automatically (based on the number of overlapping meaningful words, including their synonyms, in the OED definition and the WordNet gloss) and the results were manually validated.

The compiled dataset comprises 4,879 verb – noun **word sense pairs** with confirmed direction of derivation: 2,917 (60%) of them are zero verbs (N-to-V) and 1,962 (40%) are zero nouns (V-to-N). The bias towards zero verbs is expected, given that they are in general more productive than zero nouns. In what follows we often refer to these pairs as 'N-V pairs', by which we mean 'N-V word sense pairs'.

3 We thank James McCracken and Emily Hoyland (from the OED team) for providing us with the lists of zero nouns and zero verbs, as well as the API access.

3.4 Further remarks on methodology

In this section we briefly describe the features and the steps followed for the analysis of V-to-N and N-to-V zero derivation. Our observations are solely based on the N-V pairs derived from WordNet and labeled as morphosemantically related (cf. §3.1).

It is worth mentioning some limitations on the WordNet data:

(a) morphosemantic relations in Princeton WordNet are not comprehensive.
(b) Some morphosemantic relations are not clearly distinctive and overlap in semantics (e.g., By-means-of and Uses).
(c) Some verbs share the semantics of more than one semantic class (labeled by the semantic prime it is assigned).
(d) Some nouns share the semantics of more than one semantic class.
(e) WordNet has limited coverage.
(f) The frequency data presented here and the statistics show the distribution of the cases within the lexical-semantic network and not the frequency of their usage in text.

A further limitation occurs in the process of selecting the relevant N-V pairs:

(g) We only take those pairs for which directionality is present in the OED data (see §3.2). This excludes homonymous or other border-case examples for which the direction in the current senses of the verb and/or noun cannot be confirmed. These are the examples which are likely to exhibit non-typical features or behaviour, and introduce more variety in the data. The analysis of these examples falls among the tasks for future work.

Brief manual verification confirmed the validity of the alignment between the OED senses and the WordNet synset entries, which justifies the assignment of the direction of derivation from OED onto WordNet verb-noun pairs.

A last possible limitation concerns the reliability of the derivational direction established in the OED for zero N-V pairs. To determine this, OED lexicographers have reportedly considered the full history of each word, including date of attestation, early frequency of use, but also linguistic and etymological factors such as the behavior of cognate words, the donor in case of loanwords, and semantic properties to the extent that the more basic meaning would be associated with the base word (Philip Durkin, p. c.). These complex considerations ensure a high reliability of the derivational direction for the zero N-V pairs employed in our study (cf. directionality criteria in Plag 2003, §5.1.1.).

The results presented in §3.2 rely on simple statistical analysis of frequencies of occurrence of zero verbs and zero nouns. We apply Pearson's Chi-squared test (Manning & Schütze 1999) to check whether the direction of derivation is indepen-

dent of the morphosemantic relation, as well as of the verb and noun primes. The chi-square tests are followed by a post-hoc analysis, which observes the standardised residuals and the relative contribution of each cell to the overall chi-square statistic (Beasley & Schumacker 1995). The standardised residuals are expressed as a ratio of the difference between the observed and the expected value and the standard deviation of the expected value. They provide a measure of how significant the cells are to the chi-square value and can be used for comparing the relative contribution between cells to draw conclusions on the significance of certain classes.

In each case we first set the chi-square null hypothesis H_0 of independence and we aim at significance level of $p = 0.01$. The calculation of the chi-square values and the post-hoc analysis are carried out in `RStudio` using the `chisq.test` and the `chisq.posthoc.test` package.[4] The p-values reported in the analysis represent the probability that a certain value registered in the data can occur by chance. We look for the cells with higher standardised residuals, thus stronger contribution to the overall chi-square statistic, but also associated with low values of p corresponding to the set significant level.

The in-depth analysis of zero formations is focused on the semantic features of the verb-noun pair as expressed through the morphosemantic relations between them, as well as the verb primes and the noun primes representative for each relation. The relations are grouped based on the semantic diversity they exhibit with respect to the number of prime pairs and their compatibility. The analysis aims at providing evidence for answering the research questions proposed in §1.

4 Results

Below we first offer the overall results on the distribution of zero nouns and zero verbs for each of the 14 morphosemantic relations (§4.1), followed by an overview of the verb and noun prime distribution (§4.2). Then, in §4.3, we present the most frequent combinations of semantic primes for each relation and each direction of zero derivation, while in §4.4 we look into the possibility of having pairs of semantic primes specific to each of the two derivational directions. An interpretation of these results is offered in §5.

4 https://CRAN.R-project.org/package=chisq.posthoc.test

4.1 morphosemantic relations

An overview of the frequency of zero nouns and zero verbs for each of the 14 morphosemantic relations is given in Figure 1: in dark grey we see the numbers of zero nouns and in light grey those of zero verbs. Table 4 shows the total number of N-V word sense pairs for each relation and the corresponding percentages of zero nouns (V-to-N direction) and zero verbs (N-to-V direction) among them.

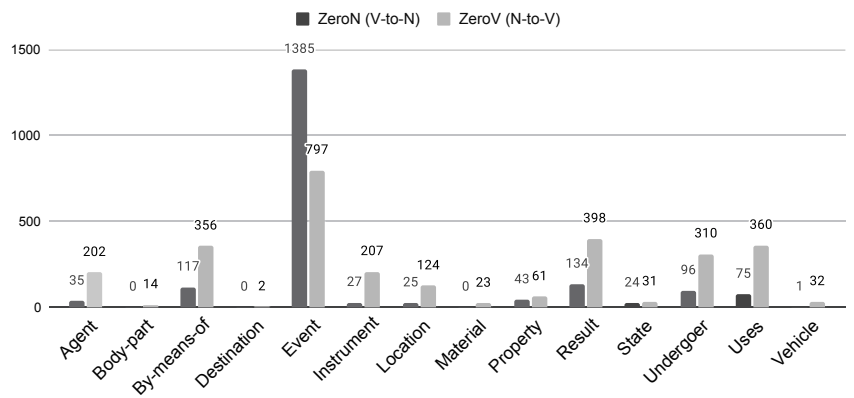


Fig. 1: Distribution of zero nouns and zero verbs per morphosemantic relation.

The Event relation is by far the most frequent with a total of 2,182 N-V pairs. In contrast, the relations Body-part, Destination, Material and Vehicle are very rare, exhibiting between 2 and 33 N-V pairs. The next most frequent relations are Result (532 N-V pairs), By-means-of (473 pairs), Uses (435 pairs), Undergoer (406 pairs), Instrument (238 pairs) and Agent (237 pairs).

The distribution of the two directions V-to-N and N-to-V within these morphosemantic relations shows some clear tendencies (see Table 4). The relations Body-part, Destination, Material and Vehicle appear (almost) exclusively with zero verbs, but they are quite infrequent in comparison with the other relations. The most frequent relation Event shows a preference for zero nouns (63.5%), while Agent, By-means-of, Instrument, Location, Result, Uses and Undergoer associate mainly with zero verbs (74.9% to 88.5%). The relations Property and State do not particularly associate with either direction: their distribution roughly corresponds to the overall distribution of zero nouns and zero verbs in the full dataset.

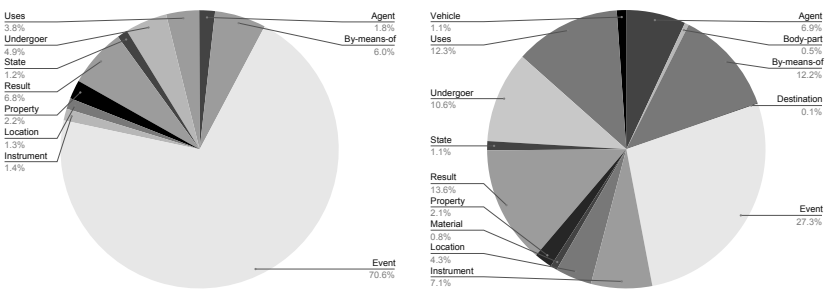
We apply the chi-square test to check the hypothesis of independence between the direction of derivation and the morphosemantic relations. The calcu-

Tab. 4: Distribution of zero nouns and zero verbs per morphosemantic relation, standardised residuals and p-values for morphosemantic relations. Light grey shading indicates relevant significant values for zero verbs and dark grey shading for zero nouns.

Relation	Agent	Body-part	By-means-of	Destination	Event	Instrument	Location
Total pairs	237	14	473	2	2,182	234	149
V-to-N	35 (14.8%)	–	117 (24.7%)	–	1,385 (63.5%)	27 (11.5%)	25 (16.8%)
N-to-V	202 (85.2%)	14 (100%)	356 (75.3%)	2 (100%)	797 (36.5%)	207 (88.5%)	124 (83.2%)
Chi^2 std. residuals ZeroN	-8.19	-3.07	-7.22	-1.16	29.80	-9.16	-5.92
Chi^2 std. residuals ZeroV	8.19	3.07	7.22	1.16	-29.80	9.16	5.92
Chi^2 p-value	0.00	0.059	1.4e-11	1.00	0.00	0.00	8.74e-08
Relation	**Material**	**Property**	**Result**	**State**	**Uses**	**Undergoer**	**Vehicle**
Total pairs	23	104	532	55	435	406	33
V-to-N	–	43 (41.3%)	134 (25.1%)	24 (43.6%)	75 (17.2%)	96 (23.6%)	1 (3%)
N-to-V	23 (100%)	61 (58.7%)	398 (74.9%)	31 (56.7%)	360 (82.8%)	310 (76.4%)	32 (97%)
Chi^2 std. residuals ZeroN	-3.94	0.24	-7.49	0.52	-7.11	-10.24	-4.37
Chi^2 std. residuals ZeroV	3.94	-0.24	7.49	-0.52	7.11	10.24	4.37
Chi^2 p-value	0.0023	1.00	2.00e-12	1.00	3.20e-11	0.00	0.00034

lated chi-square value of 953.39 ($N = 4,879$, $df = 13$, $p < 2.2e-16$) firmly rejects the hypothesis of independence.

As expected, the relations Agent, By-means-of, Instrument, Location, Result, Uses and Undergoer significantly associate with zero verbs, while Event associates with zero nouns. All these contribute significantly to the chi-value, as opposed to Property and State (among the richer relations), which show no significant association with either direction.[5]



(a) Zero nouns (total of 1,962) **(b)** Zero verbs (total of 2,917)

Fig. 2: Distribution of the morphosemantic relations within zero nouns and zero verbs.

5 Although we keep Body-Part, Destination, Material and Vehicle in the data, many researchers exclude categories where an expected value is less than 5. The inclusion of a small number of such cases is permissible and does not influence the quality of the overall results.

The distribution of the 14 relations within zero nouns and zero verbs is given in Figure 2. The Event relation contains the majority of zero nouns: i.e. a proportion of 70.6% of the 1,962 V-to-N pairs. Implicitly, all the other relations are quite infrequent with zero nouns: only By-means-of and Result come closer to the expected even distribution of 7.1%, given the 14 relations. By contrast, zero verbs exhibit a much more balanced distribution. The Event relation is still best represented, but it covers only 27.3% of the 2,917 N-to-V pairs. Six further relations such as Agent, By-means-of, Instrument, Result, Uses and Undergoer appear with around 7% or more of the N-to-V pairs each. These results show that zero nouns primarily denote events, while zero verbs are semantically more diverse, an observation we will interpret in §5.2 in light of the overview in §2.

4.2 Verb and noun prime distribution

In this section we report on how noun/verb semantic classes may associate with a derivational direction irrespective of the morphosemantic relation. These observations will be informative for the subsequent analysis of verb prime – noun prime combinations in §4.3.

We apply chi-square tests of independence for the direction of derivation and the noun and verb primes. The hypothesis of independence between the direction and the noun prime is rejected with high confidence level as the chi-square value is 937.38 ($N = 4,879$, $df = 24$, $p < 2.2e-16$). Significant contributors to the chi-value and thus, to rejecting the hypothesis of independence, are the noun primes with shaded values for the relevant direction in Table 5: noun.act, noun.animal, noun.artifact, noun.body etc.

The hypothesis of independence between the derivational direction and the verb primes is also firmly rejected although with a smaller chi-value of 189.16 ($N = 4,879$, $df = 14$, $p < 2.2e-16$). Significant contributors are the verb primes with shaded values for the relevant direction in Table 6: verb.motion, verb.perception, verb.creation and verb.possession.

The statistical data indicate strong correlations between noun and verb semantic primes and the direction of derivation. These results are discussed in the beginning of Section §5. In order to delve deeper into the correspondences between zero derivation and semantics, a next step is to look at the distribution of noun and verb prime pairs within each relation and how those are reflected in the direction of derivation. The results of this analysis are presented in §4.3.

Tab. 5: Distribution of zero nouns and zero verbs, standardised residuals and p-values for noun primes. Shading highlights the positive residuals indicating the relevant derivational direction (4th and 5th column) for the corresponding significant p-values (6th column).

	ZeroN (%)	ZeroV (%)	ZeroN std. res.	ZeroV std. res.	p-value
noun.act	722 (67.5%)	347 (32.5%)	20.62	-20.62	0.00
noun.animal	2 (4.7%)	41 (95.3%)	-4.78	4.78	8.9e−5
noun.artifact	146 (16.8%)	725 (83.2%)	-15.57	15.57	0.00
noun.attribute	114 (47.1%)	128 (52.9%)	2.24	-2.24	1.00
noun.body	10 (14.3%)	60 (85.7%)	-4.46	4.46	4.2e−4
noun.cognition	76 (38.6%)	121 (61.4%)	-0.48	0.48	1.00
noun.communication	218 (38.8%)	344 (61.2%)	-0.73	0.73	1.00
noun.event	260 (73.2%)	95 (26.8%)	13.18	-13.18	0.00
noun.feeling	24 (46.2%)	28 (53.8%)	0.88	-0.88	1.00
noun.food	23 (25.0%)	69 (75.0%)	-3.00	3.00	0.13
noun.group	22 (19.1%)	93 (80.9%)	-4.67	4.67	1.5e−4
noun.location	24 (26.1%)	68 (73.9%)	-2.79	2.79	0.26
noun.motive	0 (0.0%)	1 (100.0%)	-0.82	0.82	1.00
noun.object	30 (30.0%)	70 (70.0%)	-2.1	2.1	1.00
noun.person	34 (14.5%)	200 (85.5%)	-8.21	8.21	0.00
noun.phenomenon	27 (42.9%)	36 (57.1%)	0.43	-0.43	1.00
noun.plant	5 (10.2%)	44 (89.8%)	-4.31	4.31	8.3e−4
noun.possession	27 (30.3%)	62 (69.7%)	-1.92	1.92	1.00
noun.process	21 (56.8%)	16 (43.2%)	2.06	-2.06	1.00
noun.quantity	10 (21.7%)	36 (78.3%)	-2.57	2.57	0.52
noun.relation	2 (20.0%)	8 (80.0%)	-1.3	1.3	1.00
noun.shape	38 (40.0%)	57 (60.0%)	-0.04	0.04	1.00
noun.state	88 (46.3%)	102 (53.7%)	1.75	-1.75	1.00
noun.substance	18 (11.7%)	136 (88.3%)	-7.34	7.34	1.10e−11
noun.time	21 (41.2%)	30 (58.8%)	0.14	-0.14	1.00

4.3 Semantic classes involved in N-V pairs

In this section, we analyse the predominant pairs of semantic primes involved in both the N-to-V and the V-to-N direction for each relation.[6] This provides an overview of the semantic diversity of zero nouns and zero verbs in comparison to their bases.

For each relation and each direction, all combinations of N and V primes were extracted and filtered according to their frequency. We ignored those with poor representation and kept only the most frequent prime pairs for each direction covering around 70% of the prime pairs for each relation. The data show three types

6 We discard the relations Body-part, Destination, Material and Vehicle due to data parsimony.

Tab. 6: Distribution of zero nouns and zero verbs, standardised residuals and p-values for verb primes. Shading highlights the positive residuals indicating the relevant derivational direction (4th and 5th column) for the corresponding significant p-values (6th column).

	ZeroN (%)	ZeroV (%)	ZeroN std. res.	ZeroV std. res.	p-value
verb.body	121 (49.0%)	126 (51.0%)	2.88	-2.88	0.12
verb.change	197 (39.4%)	303 (60.6%)	-0.39	0.39	1.00
verb.cognition	61 (33.5%)	121 (66.5%)	-1.88	1.88	1.00
verb.communication	244 (40.3%)	362 (59.7%)	0.027	-0.027	1.00
verb.competition	53 (29.1%)	129 (70.9%)	-3.11	3.11	0.56
verb.consumption	35 (47.3%)	39 (52.7%)	1.25	-1.25	1.00
verb.contact	416 (38.3%)	669 (61.7%)	-1.43	1.43	1.00
verb.creation	59 (21.6%)	214 (78.4%)	-6.45	6.45	3.0e–9
verb.emotion	41 (45.6%)	49 (54.4%)	1.04	-1.04	1.00
verb.motion	329 (55.8%)	261 (44.2%)	8.22	-8.22	0.00
verb.perception	137 (58.8%)	96 (41.2%)	5.93	-5.93	9.1e–8
verb.possession	73 (26.4%)	203 (73.6%)	-4.8	4.8	4.7e–5
verb.social	98 (31.8%)	210 (68.2%)	-3.10	3.10	0.06
verb.stative	84 (43.5%)	109 (56.5%)	0.96	-0.96	1.00
verb.weather	14 (35.0%)	26 (65.0%)	-0.68	0.68	1.00

of situations with respect to the number of semantic classes (or primes) characterising the N-V pairs (see Table 7):

- the number of V primes is greater that that of the N primes, thus showing greater variety in terms of verb semantics (the dark grey cells in Table 7);
- the number of N primes is greater than that of the V primes, showing greater variety in terms of noun semantics (the white cells in Table 7);
- the number of N primes is approximately equal to that of the V primes (the light grey cells in Table 7).

Tab. 7: Overview of the comparison between the number of primes of zero formations and their bases: dark grey indicates a higher number of V primes than of N primes; white indicates a higher number of N primes; light grey indicates equal numbers of N and V primes.

	Agent	By-means-of	Event	Instrument	Location	Property	Result	State	Undergoer	Uses
N-to-V										
V-to-N										

The prime combinations yield four groups of morphosemantic relations:

- relations for which the number of V primes is higher than that of N primes, irrespective of the direction (in Table 7 both cells of these relations are in dark grey): Agent, Event, and Property (subsection §4.3.1);
- relations for which the number of N primes is higher than that of V primes, irrespective of the direction (in Table 7 both cells are white): Result and Undergoer (subsection §4.3.2);
- a relation for which the number of primes for zero verbs is higher than that of their base nouns, while for zero nouns, the number of primes of the base verbs is a little smaller than that of the primes of the zero nouns (in Table 7 the cell for N-to-V is dark grey and the one for V-to-N is white): By-means-of (subsection §4.3.3);
- relations for which the number of primes for zero verbs is higher than that of their base nouns, while for zero nouns, the number of primes of the base verbs is equal to that of the primes of the zero nouns (in Table 7 the cell for N-to-V is dark grey and the one for V-to-N is light grey): Instrument, Location, State and Uses (subsection §4.3.4).

The rest of the subsection presents the data for all the discussed relations grouped according to the tendencies outlined above. Each subsection begins with a table showing the number of N and V primes involved in each derivation direction for each morphosemantic relation and continues with a brief discussion of the data and examples of the most frequent prime pairs. A detailed representation of the N and V prime combinations for each relation is included in the appendix.

4.3.1 Verb primes prevailing for both directions

The relations Agent, Event and Property (Table 8) demonstrate a greater diversity of verb primes both as bases and as results.

Tab. 8: The number of N and V primes for the relations Agent, Event and Property.

	N-to-V		V-to-N	
	# N primes	# V primes	# V primes	# N primes
Agent	1	**9**	**6**	1
Event	11	**14**	**11**	4
Property	5	**6**	**5**	3

Agent. The relation is characterised by a single noun prime (noun.person) and many verb primes. The unique noun prime attests to the specificity of the relation. Examples of the most frequent prime pairs for N-to-V direction include:
- n.person_v.social: *clerk – to clerk*;
- n.person_v.communication: *cipher – to cipher*;
- n.person_v.creation: *cook – to cook*;
- n.person_v.competition: *jockey – to jockey*;
- n.person_v.possession: *hog – to hog*;
- n.person_v.cognition: *pioneer – to pioneer*.

Some examples of the most frequent prime pairs for V-to-N direction include:
- v.social_n.person: *to affiliate – affiliate*;
- v.communication_n.person: *to flirt – flirt*;
- v.motion_n.person: *to sneak – sneak*.

Event. This relation is characterised by a diversity of verb and noun semantic primes. Slightly more semantic verb classes are involved when zero verbs are created than in the case of zero nouns (see 14 vs. 11 in Table 8). Examples for the most represented combinations of primes for zero verbs include:
- n.communication_v.communication: *cheer – to cheer*;
- n.act_v.motion: *curvet – to curvet*;
- n.act_v.contact: *contact – to contact*;
- n.act_v.social: *crusade – to crusade*;
- n.act_v.possession: *finance – to finance*;
- n.event_v.motion: *progress – to progress*.

The most frequent combinations of primes for creating zero noun Events include:
- v.motion_n.act: *to prowl – prowl*;
- v.motion_n.event: *to quake – quake*;
- v.contact_n.act: *to pull – pull*;
- v.communication_n.communication: *to boast – boast*;
- v.change_n.act: *to purge – purge*.

A very neat distribution is observed with respect to the noun primes involved in the relation: although the base noun and the zero verb classes are quite diverse (11 noun primes and 14 verb primes), the most frequent noun primes (both in terms of the number of pairs and the diversity of the primes of the resulting verbs) serving as bases for verbs from various primes are in fact only two: n.act and n.event (see the shaded rows in the N-to-V part of Table 14 in the appendix). The same holds in the opposite direction: 11 verb primes are bases for only 4 noun primes, of which

noun.act and noun.event are by far the most frequent (the shaded columns in the V-to-N part of Table 14).

Property. At first glance the data on N-to-V derivation with this relation do not show greater variety of verb primes: the resulting verbs of the 5 base noun primes belong to 6 verb primes; however, the majority of derivations and the greatest diversity of resulting verb primes are concentrated in the most productive noun prime (noun.attribute) (the shaded row in the N-to-V part of Table 15 in the appendix). Here are several examples to illustrate the most frequent N-V pairs:

- n.attribute_v.communication: *glamour – to glamour*;
- n.attribute_v.change: *silence – to silence*;
- n.attribute_v.cognition: *order – to order*;
- n.attribute_v.contact: *grade – to grade*;
- n.attribute_v.social: *disadvantage – to disadvantage*.

As compared with the base verbs (5 primes), zero nouns involve a more restricted number of primes (3). At a closer look, we also find that various base verb primes (the shaded column in the V-to-N part of Table 15 in the appendix) result in semantically uniform nouns denoting properties and attributes (noun.attribute). The only exception with respect to the noun primes is verb.stative, which results in location and relation nouns, too:

- v.perception_n.attribute: *to feel – feel*;
- v.contact_n.attribute: *to polish – polish*;
- v.stative_n.attribute: *to overlap – overlap*;
- v.stative_n.location: *to reach – reach*;
- v.stative_n.relation: *to trim – trim*;
- v.motion_n.attribute: *to slope – slope*;
- v.change_n.attribute: *to simmer – simmer*.

In addition, the diversity of verb primes with Property is much more limited than the one attested with relations such as Event or Agent. This, along with the fact that attributes are both the best represented bases and results, points to the conclusion that this relation is more specialised than, for instance, Event.

4.3.2 Noun primes prevailing for both directions

The relations Result and Undergoer (Table 9) are characterised by a greater diversity of noun primes both as bases and as results of the derivation. We should note, however, that base nouns are only slightly more diverse than the resulting zero verbs.

Tab. 9: The number of N and V primes for the relations Result and Undergoer.

	N-to-V		V-to-N	
	# N primes	# V primes	# V primes	# N primes
Result	**13**	11	7	**13**
Undergoer	**12**	11	10	**13**

Result. At first sight, the number of base noun primes (13) exceeds the number of resulting verb primes (11) by only a little. However, if we consider the variety of base noun primes in light of the concentration of the resulting verb primes, it is obvious that 3 of the latter, verb.contact, verb.creation and verb.change, are more frequent (as shown by the shaded cells in Table 16 in the appendix). Below are examples of the most frequent N-to-V prime pairs:

- n.artifact_v.contact: *knot – to knot*;
- n.shape_v.contact: *groove – to groove*;
- n.object_v.contact: *segment – to segment*;
- n.artifact_v.creation: *fresco – to fresco*;
- n.communication_v.creation: *film – to film*;
- n.object_v.change: *splinter – to splinter*;
- n.food_v.change: *stew – to stew*;
- n.attribute_v.change: *bleach – to bleach*;
- n.substance_v.change: *rust – to rust*.

With respect to the zero nouns, their number of primes (13) is substantially greater than that of the verbs (7).[7] The best represented base verb primes, v.contact, v.creation and v.change, are the same as the resulting verb primes in the N-to-V derivation (see the shaded cells in Table 16 in the appendix). Below are given a few examples of the patterns found in the data:

- v.contact_n.shape: *to scratch – scratch*;
- v.contact_n.state: *to scrape – scrape*;
- v.contact_n.object: *to slice – slice*;
- v.change_n.shape: *to crease – crease*;
- v.change_n.attribute: *to tinge – tinge*;
- v.change_n.object: *to split – split*;
- v.creation_n.artifact: *to knit – knit*;
- v.creation_n.communication: *to reissue – reissue*.

7 The data for the relation Result are estimated for 80-81% of the prime pairs (not 70% as for the others) because of the large number of pairs with the same number of attestations.

The productivity of verb.contact, verb.change and verb.creation as both bases and results of the derivation resonates well with the semantics of the relation. In the V-to-N direction, change verbs and creation verbs are naturally related to results denoting the entity that is formed or undergoes a change (*to curl – a curl*) or comes into existence (*to produce – produce*) or is affected (*to scrape – a scrape*) as a result of the verb's event. Many of the base nouns in the N-to-V direction are concrete entities, with primes such as noun.object, noun.artifact, noun.food, noun.substance.

Undergoer. For the N-to-V direction the number of noun primes involved exceeds the verb primes only by a little. Although we find 12 base noun primes, only two of them, i.e. noun.communication and noun.artifact, produce verbs from various primes (see the first 2 shaded rows in the N-to-V part of Table 17 in the appendix). The remaining ones yield verbs that are concentrated in a limited number of primes, most favoured being verb.contact and verb.change (see the shaded columns in the N-to-V part of Table 17). Here are some examples of the most frequent prime pairs:

- n.communication_v.communication: *decree – to decree*;
- n.communication_v.creation: *solo – to solo* '= perform a solo';
- n.artifact_v.contact: *freight – to freight*;
- n.artifact_v.creation: *microfilm – to microfilm*;
- n.possession_v.possession: *allowance – to allowance*.

In the V-to-N direction we find 10 base verb primes and 13 zero noun primes. Even so, we should note that there is a marked tendency for verbs from diverse primes to form nouns denoting artifacts (the shaded column in the V-to-N direction in Table 17), with most of the verb primes producing nouns from only 3 primes. The resulting nouns are quite scattered across primes although there is logic to this distribution: verbs of communication form nouns of communication, verbs of consumption – nouns denoting food, etc. This goes in line with the default assumption that the resulting zero nouns are the natural affected entities (undergoers) of the respective predicates:

- v.possession_n.possession: *to refund – refund*;
- v.contact_n.artifact: *to cover – cover*;
- v.communication_n.communication: *to rehash – rehash*;
- v.consumption_n.food: *to feed – feed*;
- v.body_n.body: *to slobber – slobber*.

4.3.3 Zero verbs – semantically more diverse than their base nouns; zero nouns – semantically slightly more diverse than their base verbs

Tab. 10: The number of N and V primes for the relation By-means-of.

	N-to-V		V-to-N	
	# N primes	# V primes	# V primes	# N primes
By-means-of	8	11	10	11

By-means-of. This relation is characterised by great diversity with respect to the semantic classes to which the verbs and nouns involved (either as bases or as results of zero derivation) belong. Table 10 shows that 8 noun primes are bases for 11 zero verb primes. However, a closer inspection reveals that two semantic classes of nouns (n.artifact and n.communication) are bases for a wide spectrum of verb classes (see the shaded lines in Table 18 in the appendix). Some examples from the most representative pairs of primes include:

– n.artifact_v.contact: *border* – *to border*;
– n.artifact_v.motion: *fan* – *to fan*;
– n.communication_v.communication: *certificate* – *to certificate*.

For zero nouns, although 10 verb primes are involved as bases for deriving nouns belonging to 11 primes, most of the created nouns tend to belong to the semantic classes noun.artifact and noun.attribute, while their base verbs cover a wide semantic range for this direction of derivation too (see the shaded columns in Table 18). Some examples of the prevailing primes combinations are:

– v.contact_n.artifact: *to cover* – *cover*;
– v.social_n.attribute: *to control* – *control*.

Tab. 11: The number of N and V primes for the relations Instrument, Location, State and Uses.

	N-to-V		V-to-N	
	# N primes	# V primes	# V primes	# N primes
Instrument	1	5	1	1
Location	3	6	3	3
State	4	6	3	3
Uses	8	9	6	6

4.3.4 Zero verbs – semantically more diverse than their base nouns; zero nouns – as semantically diverse as their base verbs

Instrument. Just like Agent, this relation involves only one noun prime, i.e. noun.artifact, irrespective of the derivation direction, and various verb semantic classes (see Table 19 in the appendix). Examples of these combinations are:

- n.artifact_v.contact: *hammer – to hammer*;
- n.artifact_v.creation: *pencil – to pencil*;
- n.artifact_v.motion: *pedal – to pedal*;
- n.artifact_v.body: *splint – to splint*;
- n.artifact_v.change: *hose – to hose*.

For zero nouns only one combination of primes occurs for the majority of pairs: v.contact – n.artifact (e.g., *lift – to lift*).

Location. The semantic diversity of zero verbs is greater than that of the base nouns. Some of the most representative combinations are:

- n.artifact_v.contact: *barrel – to barrel*;
- n.artifact_v.motion: *dock – to dock*;
- n.location_v.motion: *corner – to corner*;
- n.object_v.motion: *puddle – to puddle*.

The zero nouns of type Location display only a few pairs of semantic classes for 3 verb and 3 noun primes (see Table 20 in the appendix). We give below examples for these combinations:

- v.contact_n.location: *to scour – scour*;
- v.change_n.attribute: *to burn – burn*;
- v.motion_n.location: *to stop – stop*;
- v.motion_n.artifact: *to walk – walk*.

State. The N-to-V direction of derivation for State shows 4 noun primes and 6 verb primes. The prime noun.state is best represented, and the rest – noun.feeling, noun.quantity and noun.attribute – are clearly stative in nature, too. The resulting verbs belong to various primes (see Table 21 in the appendix). Below are some examples of the prime pairs:

- n.state_v.contact: *contact – to contact*;
- n.state_v.emotion: *miff – to miff*;
- n.state_v.change: *mess – to mess*;
- n.feeling_v.emotion: *pity – to pity*;

The zero nouns involve 3 base verb primes and 3 resulting noun primes. However, as the data further show (Table 21), most of the zero nouns are states and the remaining ones have stative semantics, like for zero verbs. As the number of prime pairs considered is small, below we list examples for all of them:

- v.emotion_n.state: *to glow – glow*;
- v.emotion_n.feeling: *to dread – dread*;
- v.change_n.state: *to polish – polish*;
- v.change_n.attribute: *to temper – temper*;
- v.body_n.state: *to frazzle – frazzle*.

Uses. In the N-to-V direction of derivation for this relation there are 8 noun primes and 9 verb primes. The most productive base prime, noun.artifact, produces verbs from various primes (see Table 22 in the appendix). Nouns of the classes noun.substance and noun.communication, which are also productive, are bases for verbs belonging to the most numerous semantic primes occurring with this relation.

The remaining noun primes yield a modest number of verbs from different primes. In fact, we note that nouns from various primes tend to produce verbs primarily with the primes verb.contact and verb.possession. Examples of the most productive primes pairs are:

- n.artifact_v.contact: *canopy – to canopy*;
- n.artifact_v.possession: *armour – to armour* 'equip with armour';
- n.artifact_v.body: *powder – to powder*;
- n.substance_v.contact: *copper – to copper*;
- n.communication_v.communication: *autograph – to autograph*.

The V-to-N direction of derivation involves an even number of 6 primes for both nouns and verbs; in addition, the most productive prime verb.contact yields nouns from various primes, with the higher number of noun.artifact,

n.communication and n.substance (see Table 22). Below are examples of the productive patterns:

- v.contact_n.artifact: *to paint – paint*;
- v.contact_n.communication: *to stamp – stamp*;
- v.contact_n.substance: *to daub – daub*;
- v.body_n.artifact: *to attire – attire*;
- v.creation_n.artifact: *to trim – trim*.

4.4 Prime combinations and the direction of zero derivation

In this section we consider the possibility of having prime pairs that manifest the tendency of occurring with zero derivations of a certain direction.

The 25 noun primes and the 15 verb primes are found in 263 combinations in our dataset, 40 of which are hapax legomena and 23 have more than 50 occurrences. We present here the results of the investigation we made for establishing their specificity to a certain direction of zero derivation.[8] The first remark is that most prime combinations (163 pairs) are found with both directions, 24 are specific to zero nouns formation, while 76 are specific to zero verbs formation.

From the prime combinations that are specific to one direction of derivation, we highlight here the more frequent ones: for zero nouns, only one combination, namely v.stative – n.event (11 occurrences) is specific, while several pairs are specific to zero verb formation:

- v.creation – n.person (17 occurrences) – specific to Agent only,
- v.contact – n.animal (16 occurrences),
- v.contact – n.group (15 occurrences),
- v.competition – n.person (14 occurrences) – specific to Agent only,
- v.competition – n.group (12 occurrences),
- v.body – n.person (12 occurrences),
- v.contact – n.plant (11 occurrences),
- v.body – n.animal (11 occurrences),
- v.change – n.cognition (10 occurrences),
- v.perception – n.person (10 occurrences) – specific to Agent only.

We subjected to a closer inspection 23 pairs of primes combinations that (i) occur more than 50 times in the dataset and (ii) display a frequency of occurrence for a

8 There is only one example of zero derivation for each of the hapax legomena in the dataset.

derivational direction higher than expected.[9] The data showed that some of these prime pairs are specific to a certain direction when occurring with a certain morphosemantic relation. This happens mostly with zero verbs of creation, as shown in Table 12.[10]

Tab. 12: Semantic primes specific to certain relations for zero verbs creation: the number shows how many zero derivations exist for the respective primes combination and relation; frequency is calculated with respect to the whole number of derivations of the respective direction for all relations.

Primes pair	Relation	#occur.	Freq.	Examples
n.artifact_v.creation	Instrument	12	14%	*chalk*$_N$ - *chalk*$_V$, *charcoal*$_N$ - *charcoal*$_V$, *pencil*$_N$ - *pencil*$_V$
n.artifact_v.creation	Undergoer	11	13%	*bead*$_N$ - *bead*$_V$, *drum*$_N$ - *drum*$_V$, *microfilm*$_N$ - *microfilm*$_V$
n.substance_v.contact	By-means-of	9	12%	*cement*$_N$ - *cement*$_V$, *glue*$_N$ - *glue*$_V$, *grout*$_N$ - *grout*$_V$
n.substance_v.contact	Material	10	13%	*copper*$_N$ - *copper*$_V$, *metal*$_N$ - *metal*$_V$, *silver*$_N$ - *silver*$_V$
n.artifact_v.possession	Location	9	18%	*bank*$_N$ - *bank*$_V$, *garage*$_N$ - *garage*$_V$, *shop*$_N$ - *shop*$_V$
n.communication_v.creation	Undergoer	11	26%	*madrigal*$_N$ - *madrigal*$_V$, *hymn*$_N$ - *hymn*$_V$, *paragraph*$_N$ - *paragraph*$_V$

5 Interpretation of the results

We now proceed to interpret the results presented above from the perspective of our initial research questions and the literature overview in §2. Our first aim was to check whether the 14 morphosemantic relations annotated in the PWN standoff file and the PWN noun/verb primes show any correlations with the direction of derivation in our dataset of N-V word sense pairs. The second aim was to interpret these findings in the context of the previous literature.

9 The expected frequency is the representation of V-to-N and N-to-V forms in the overall dataset, i.e., 40% and 60%, respectively (see §3.3). The thresholds we established, higher than these frequencies, are 60% for zero nouns and 75% for zero verbs.

10 We left aside those cases when the occurrences represent less than 10% of the whole number of derivations for the respective direction.

To begin with, the results in §4.1 and §4.2 allow us to state that we do find strong correlations with the derivational direction both for the morphosemantic relations and the noun/verb primes. Table 4 in §4.1 reveals statistically significant correlations for most of the morphosemantic relations and one direction of derivation, which we discuss in more detail in §5.1 below.
Section §4.2 (see Tables 5 and 6) shows that 9 of the 25 noun primes and 4 of the 15 verb primes establish statistically significant correlations with one direction or the other. Nouns of the classes noun.act and noun.event associate with zero nouns, while noun.animal, noun.artifact, noun.body, noun.group, noun.person, noun.plant and noun.substance associate with zero verbs. Ontologically speaking, these results conform to our expectations about lexical categories. Verbs ontologically denote acts and events, while nouns mostly refer to objects (Baker & Croft 2017). Consequently, nouns that denote acts and events will be the output of deverbal derivation (see *transfer*, *escape*, *ambush*, *beat* for acts and *twinkle*, *twist*, *drip* for events derived from verbs), while nouns referring to animals, artifacts, persons, plants, substances will represent the input to denominal verbs (see *lamb*, *cub*, *kitten* for noun.animal, *bandage*, *chain*, *buckle* for noun.artifact, *chair*, *champion*, *doctor* for noun.person, *fruit*, *mushroom*, *seed* for noun.plant and *cement*, *glue*, *metal*, *paper* for noun.substance as input to zero verbs).

For the verb primes in Table 6, we see that verb.motion and verb.perception associate with zero nouns (see *to march*, *to limp*, *to ride*, *to sail* for verb.motion and *to smell*, *to sound*, *to scent*, *to knock* for verb.perception, as input to zero nouns), while verb.creation and verb.possession associate with zero verbs (see *to hammer*, *to ornament*, *to instrument*, *to garden* for verb.creation and *to loan*, *to profit*, *to sacrifice*, *to fund* for verb.possession, as zero verbs derived from nouns). This suggests that verbs of motion and of perception are more likely to form the input to zero nouns than to be derived from lexical nouns, while verbs of creation and verbs of possession are more likely derived from nouns.

5.1 Interpretation of the morphosemantic relations

In §4.1, we saw that the Event relation is the most frequent one in the dataset and substantially more frequent with zero nouns (63.5%) than with zero verbs (36.5%), if we keep in mind that the overall dataset is biased towards zero verbs (60%). The relation Event between nouns and verbs is expected to be frequent, given the primarily eventive meaning of verbs (Koontz-Garboden 2005; Baker & Croft 2017). Its higher frequency with zero nouns, however, also confirms the intuition that, to create an Event relation, the input category must be that of the verb, whose primary ontological meaning is eventive. The noun is the output category, as events

are not typical denotations in the ontology of nouns (pace some exceptions like *trip*, *movie*).

Unlike Event, relations such as Agent, By-means-of, Instrument, Location, Result, Undergoer and Uses are substantially more frequent with zero verbs, which represent 75% or more of the corresponding pairs.[11] These relations are relatively frequent in our dataset and represent typical morphosemantic relations that appear between morphologically related N-V pairs, to the extent that the noun loosely corresponds to some semantic argument in the event structure of the related verb. From this perspective, it is natural for them to appear more frequently with zero verbs, as the meaning of the verb is construed around a semantic argument instantiated by the base noun (but see §2.2 for more subtle observations on this).

An interesting case is that of the Property and State relations, which are not very frequent in our dataset, nor do they show any tendency towards one direction or the other. From the literature on the semantics of lexical categories we know that in English properties and states are typically categorised as adjectives, and not as verbs or nouns (Dixon 1982; Koontz-Garboden 2005, 2007; Baker & Croft 2017; Koontz-Garboden & Francez 2017). This explains the smaller number of N-V pairs for these relations to begin with but also their indeterminacy with respect to a derivational direction, since neither the verb nor the noun represents a default base category for zero formations that denote properties and states (in contrast with Event, Agent, Instrument, etc. above).

From the perspective of how zero nouns and zero verbs are distributed across the different relations in Figure 2, §4.1, we can also make some insightful observations, which support previous claims in the literature. Namely, the figure shows that zero nouns are primarily formed to denote events: 70.3% of them appear in this relation, indicating a clear morphosemantic relation to the base from which they inherit an eventive meaning. Other relations are much more rarely attested with zero nouns: only By-means-of and Result come close to what would be expected from a by-chance distribution (i.e. 7.1%). This picture confirms Cetnarowska's (1993) conclusion that zero nouns primarily denote events, and participant readings are the result of metonymic shifts and only limitedly available. For zero verb formation, however, we find a broader spectrum of morphosemantic relations, indicating more indeterminacy about their semantics in relation to the base: Agent, By-means-of, Event, Instrument, Result, Undergoer and Uses include between 6.9% and 27.3% of the zero verbs. This contrast be-

11 The dataset is indeed biased for zero verbs, which make 60% of it, but these relations show a considerably higher proportion than this baseline (see statistical results in §4.1).

tween zero nouns and zero verbs supports the observation from the literature overview in §2 and the results from Kisselew et al. (2016) according to which the semantic relationship between zero nouns and their base is more systematic and predictable with information content measures than that between zero verbs and their base.

5.2 Interpretation of the noun and verb prime distribution

Zero nouns have been argued to be well-behaved and exhibit meanings similar to those of overt nominalisations (Cetnarowska 1993), which are semantically more specific and restricted than the meaning of the base (Kisselew et al. 2016). Zero verbs, however, have been argued not to receive more specific meanings than the base (Kisselew et al. 2016) and to show a broader range of meanings than overt derivations, whereby they do not establish systematic relations to the meaning of their base as in an argument-predicate relation (see Harley & Haugen 2007; Rimell 2012). The question is whether our dataset allows us to draw any conclusions with respect to these observations.

More evidence on the greater semantic diversity of zero verbs compared to zero nouns in support of these previous findings appears when we interpret the results from §4.3 on the prime pairs (i.e. semantic verb and noun classes) involved in each such relation. For many relations, a small set of base noun primes results in verbs of many more primes, while in zero noun derivation the opposite holds: a broad range of base verb primes yields nouns grouped in only 1 or 2 noun primes. This contrast is particularly surprising, if we take into account that the pool of noun primes is much larger than that of verb primes in our dataset (see 25 noun primes vs. 15 verb primes in §3.1 and §4.2) and implicitly provides stronger support for the previous observation that zero verbs are semantically more diverse (in relation to their base and by comparison to overtly derived verbs) than zero nouns are.

Looking closely at this contrast, we notice that for the N-to-V direction, the extent of the phenomenon is quite vast (see Table 7), manifesting for 8 out of the 10 relations analysed. For zero nouns it is more moderate: three relations (Agent, Event and Property) clearly manifest this, two (By-means-of and State) manifest it for the most frequent prime combinations, while others (Result and Undergoer) seem not to confirm it, and still others (Instrument, Location and Uses) contain insufficient data for drawing any conclusion.

With respect to the correlation between prime combinations and the direction of zero derivation (§4.4), we noticed that most of the primes combinations occur

with both directions, so there are only a few such correlations, and most of them concern zero verbs, which is again suggestive of their more diverse semantics.

6 Conclusions and future work

In this paper we investigated whether morphosemantic relations and noun/verb semantic classes (primes) from the PWN show any correlations with the OED direction of derivation in a dataset of 4,879 zero-related N-V word sense pairs. We aimed to capture possible semantic differences between nouns and verbs formed by zero derivation for a better understanding of this morphological process.

We found statistically significant correlations with the derivational direction both for the morphosemantic relations and the semantic classes. First, we have shown that most of the morphosemantic relations associate with one direction or the other: Event appears most often with zero nouns created from verbs, while Agent, By-means-of, Instrument, Location, Result, Undergoer, Uses appear most often with zero verbs created from nouns. The relations Property and State appear with similar proportions of zero nouns and zero verbs. We explained this peculiar behaviour by the fact that properties and states are unusual denotations for verbs and nouns in English. Consequently, we expect them to show no particular tendency towards one direction or the other, and we also expect many noun-verb pairs belonging to this relation to be undetermined in terms of derivational direction. The latter is something to investigate in future research on the basis of zero pairs whose directionality is not known.

Second, we have shown that, for most relations, irrespective of the derivational direction, zero verbs are characterised by greater diversity than zero nouns. This means that within the relations we discussed, zero nouns usually show fewer or just as many primes as the corresponding base verbs, while zero verbs often show more primes than the corresponding base nouns, despite the fact that there are more noun primes than verb primes in the PWN ontology. This picture supports observations from the previous literature, according to which zero nouns are semantically more restricted in relation to the base than zero verbs are.

One important aspect of our dataset is that it contains N-V word *sense* pairs. At this point, the OED does not provide a derivational direction for individual senses. However, Plank (2010) convincingly argues that zero-derived pairs of polysemous words may show both derivational directions for different senses. In English the word *taxi* offers such an example: the verb *taxi* with the sense 'to travel in a taxi' is derived from the noun with the sense 'taxicab', but the noun with the sense 'an act or instance of taxiing' is derived from the zero verb's new sense '(of an

aeroplane) to travel slowly along the ground before take-off or after landing'. Our current dataset includes only the senses available with the N-to-V derivation (the zero verb), but we believe that a larger study of this kind would be able to distinguish between different senses associated with the same lexeme but with different derivational directions.

Zero derivation affects not only the lexical category, but also the meaning of new formations. Our employment of general lexical resources with large sets of data (the OED and the WordNet) to investigate semantic parameters available in N-V zero pairs promises a better understanding of whether the formation of zero nouns follows similar rules as the formation of zero verbs. This has further implications for their modeling and for a more refined view on this word formation process in general.

A possible application of the results obtained (and to be further elaborated) is the implementation of linguistically-informed computational models for predicting the direction of derivation for pairs of senses for which such information is not available yet, whether they prove to be directional or not (see Darby 2015; Darby & Lahiri 2016). Our dataset is a good point of departure for developing methods to predict the morphosemantic relation of previously unattested pairs, based on the distribution of the derivational direction across semantic primes and morphosemantic relations observed here. The research presented here also opens the way to similar endeavours for languages other than English, provided that similar lexical resources are available.

Acknowledgments

We thank the two reviewers and the editors for their thought-provoking feedback to previous versions of this paper. Iordăchioaia's contribution was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 404208593.

Bibliography

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

Baker, Mark & William Croft. 2017. Lexical categories: Legacy, lacuna, and opportunity for functionalists and formalists. *The Annual Review of Linguistics* 3. 179–197.

Beasley, Mark & Randall Schumacker. 1995. Multiple regression approach to analyzing contingency tables: Post hoc and planned comparison procedures. *Journal of Experimental Education* 64. 79–93.

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. 2015. *The processing of conversion in English: Morphological complexity and underspecification*. Oxford: University of Oxford dissertation.

Darby, Jeannique & Aditi Lahiri. 2016. Covert morphological structure and the processing of zero-derived words. *The Mental Lexicon* 11.2. 186–215.

Dixon, Robert M. W. 1982. *Where have all the adjectives gone?* Berlin: De Gruyter.

Fellbaum, Christiane (ed.). 1998. *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.

Fellbaum, Christiane, Anne Osherson & Peter E. Clark. 2009. Putting semantics into WordNet's 'morphosemantic' links. In Z. Vetulani & H. Uszkoreit (eds.), *Human language technology. challenges of the information society. ltc 2007*, 350–358. Berlin, Heidelberg: Springer.

Fillmore, Charles J. 1968. The case for case. In Emmon Bach & Robert T. Harms (eds.), *Universals in linguistic theory*, 1–88. New York: Holt, Rinehart, and Winston.

Gildea, Daniel & Daniel Jurafsky. 2000. Automatic labeling of semantic roles. In Anna Braasch & Claus Povlsen (eds.), *Proceedings of the ACL 2000*, 512–520. Hong Kong: Association for Computational Linguistics.

Hale, Ken & Jay Keyser. 2002. *Prolegomenon to a theory of argument structure*. Cambridge, MA: MIT Press.

Harley, Heidi & Jason D. Haugen. 2007. Are there really two different classes of instrumental denominal verbs in english? *Snippets* 16. 9–10.

Huyghe, Richard, Alizée Lombard, Justine Salvadori & Sandra Schwab. 2023. Semantic rivalry between French deverbal neologisms in *-age*, *-ion* and *-ment*. In Sven Kotowski & Ingo Plag (eds.), *The semantics of derivational morphology. Theory, methods, evidence*, 143–175. Berlin/Boston: De Gruyter.

Iordăchioaia, Gianina. 2020. Categorization and nominalization in zero nominals. In Artemis Alexiadou & Hagit Borer (eds.), *Nominalization: 50 Years on from Chomsky's Remarks*, 231–253. Oxford: Oxford University Press.

Irmer, Roland. 1972. *Die mit Nullmorphem abgeleiteten deverbalen Substantive des heutigen Englisch*. Tübingen: University of Tübingen dissertation.

Kastovsky, Dieter. 1968. *Old English deverbal substantives derived by means of a zero morpheme*. Esslingen am Neckar: Langer.

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.

Kisselew, Max, Laura Rimell, Alexis Palmer & Sebastian Padó. 2016. Predicting the direction of derivation in English conversion. In *Proceedings of the ACL SIG-MORPHON workshop*, 93–98. Berlin: Association for Computational Linguistics. https://aclanthology.org/W16-2015.

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. https://aclanthology.org/2016.gwc-1.26.

Koontz-Garboden, Andrew. 2005. On the typology of state/change of state alternations. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 2005*, 83–117. Dordrecht: Springer.

Koontz-Garboden, Andrew. 2007. *States, changes of state, and the monotonicity hypothesis*. Stanford, CA: Stanford University dissertation.

Koontz-Garboden, Andrew & Itamar Francez. 2017. *Semantic and morphosyntactic variation: Qualities and the grammar of property concepts*. Oxford: Oxford University Press.

Lieber, Rochelle. 1981. *On the organization of the lexicon*. Bloomington, OH: Indiana University Linguistics Club.

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.

Lieber, Rochelle. 2016. *English nouns. The ecology of nominalization*. Cambridge: Cambridge University Press.

Lieber, Rochelle. 2023. Ghost aspect and double plurality. In Sven Kotowski & Ingo Plag (eds.), *The semantics of derivational morphology. Theory, methods, evidence*, 15–35. Berlin/Boston: De Gruyter.

Manning, Chris & Hinrich Schütze. 1999. *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.

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

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.

Mititelu, Verginica Barbu, 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.

Padó, Sebastian, Alexis Palmer, Max Kisselew & Jan Šnajder. 2015. Measuring semantic content to assess asymmetry in derivation. In *Proceedings of the IWCS workshop on advances in distributional semantics*, London. https://www.nlpado.de/~sebastian/pub/papers/iwcs15_pado.pdf.

Plag, Ingo. 1999. *Morphological productivity: Structural constraints in English derivation*. Berlin and New York: De Gruyter.

Plag, Ingo. 2003. *Word-formation in English*. Cambridge: Cambridge University Press.

Plag, Ingo, Lea Kawaletz, Sabine Arndt-Lappe & Rochelle Lieber. 2023. Analogical modeling of derivational semantics. Two case studies. In Sven Kotowski & Ingo Plag (eds.), *The semantics of derivational morphology. Theory, methods, evidence*, 103–141. Berlin/Boston: De Gruyter.

Plank, Frans. 2010. Variable direction in zero-derivation and the unity of polysemous lexical items. *Word Structure* 3.1. 82–97.

Rimell, Laura. 2012. *Nominal roots as event predicates in English denominal conversion verbs*: University of New York dissertation.

Ruppenhofer, Josef, Colin F. Baker & Charles J. Fillmore. 2002. The FrameNet database and software tools. In Anna Braasch & Claus Povlsen (eds.), *Proceedings of the tenth EURALEX international congress*, 371–375. Copenhagen: EURALEX.

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.

Valera, Salvador. 2014. Conversion. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of derivational morphology*, 154–168. Oxford: Oxford University Press.

Appendix

Below we present a detailed overview of the N and V prime combinations for each relation and direction: the numbers and percentages are calculated with respect to the total number of N-V pairs across the observed pairs of prime combination within each relation and direction.

Tab. 13: The combination of semantic primes to express the relation Agent. The grey cells show that only one noun prime (n.person) combines with various verb primes in the data.

N-to-V	v.social	v.communication	v.creation	v.competition	v.possession	v.cognition	v.body	v.state	v.motion
n.person	53 (26%)	20 (10%)	17 (8%)	14 (7%)	13 (6%)	12 (6%)	11 (5%)	11 (5%)	11 (5%)
V-to-N	**n.person**								
v.social	7 (20%)								
v.commun	6 (17%)								
v.motion	4 (11%)								
v.poss	2 (6%)								
v.consum	2 (6%)								
v.emotion	2 (6%)								

Tab. 14: The combination of semantic primes to express the relation Event. The grey cells highlight the prime combinations in which 2 noun primes (n.act and n.event) are paired with a larger variety of verb primes.

N-to-V	v.communication	v.motion	v.contact	v.social	v.possession	v.change	v.creation	v.body	v.perception	v.emotion	v.cognition	v.weather	v.consumption	v.stative
n.communication	88 (11%)						8 (1%)				8 (1%)			
n.act	23 (3%)	52 (7%)	51 (6%)	44 (6%)	33 (4%)	23 (3%)	19 (2%)	17 (2%)	5 (1%)		11 (1%)			6 (1%)
n.event	8 (1%)	22 (3%)	9 (1%)			14 (2%)			17 (2%)					
n.feeling										17 (2%)				
n.state				7 (1%)		7 (1%)		13 (2%)						
n.cognition				7 (1%)					10 (1%)		9 (1%)			
n.possession					10 (1%)									
n.process						8 (1%)								
n.phenomenon						6 (1%)						7 (1%)		
n.food													6 (1%)	
n.time						6 (1%)								

V-to-N	n.communication	n.event	n.act	n.cognition
v.contact		39 (3%)	184 (13%)	
v.motion		53 (4%)	194 (14%)	
v.communication	106 (8%)	41 (3%)	39 (3%)	
v.change		36 (3%)	55 (4%)	
v.body	23 (2%)		48 (3%)	
v.perception		43 (3%)	22 (2%)	16 (1%)
v.social			42 (3%)	
v.competition			36 (3%)	
v.stative			26 (2%)	
v.possession			26 (2%)	
v.cognition				19 (1%)

Tab. 15: The combination of semantic primes to express the relation Property. The grey cells show that only one noun prime (n.attribute) can be the base for more verb primes in the N-to-V direction and different verb primes are base to create zero nominalisations of mostly one semantic class (also n.attribute).

N-to-V	v.communication	v.change	v.cognition	v.contact	v.social	v.emotion
n.attribute	9 (15%)	8 (13%)	6 (10%)	3 (5%)	3 (5%)	
n.communication	4 (7%)					
n.cognition			4 (7%)			2 (3%)
n.state			2(3%)			
n.quantity	2 (3%)					

V-to-N	n.attribute	n.location	n.relation
v.perception	7 (16%)		
v.contact	7 (16%)		
v.stative	3 (7%)	5 (12%)	2 (5%)
v.motion	5 (12%)		
v.change	3 (7%)		

Tab. 16: The combination of semantic primes to express the relation Result. The grey cells show the three verb primes (v.contact, v.creation and v.change) are combined with more diverse noun primes to account for most of the data.

N-to-V	v.contact	v.creation	v.change	v.communication	v.motion	v.body	v.cognition	v.competition	v.possession	v.social	v.stative		
n.artifact	33 (8%)	32 (8%)	5 (1%)	6 (1.5%)	3 (1%)								
n.object	11 (3%)	3 (1%)	17 (4%)										
n.communication		15 (4%)		14 (3.5%)									
n.shape	14 (3.5%)	3 (1%)	9 (2%)		5 (1%)						3 (1%)		
n.food	6 (1.5%)		11 (3%)										
n.attribute	3 (1%)	4 (1%)	10 (2.5%)										
n.substance	3 (1%)		10 (2.5%)			3 (1%)							
n.group	8 (2%)		5 (1%)		9 (2%)		5 (1%)	5 (1%)		3 (1%)			
n.plant			7 (2%)			3 (1%)							
n.animal						7 (2%)							
n.cognition		5 (1%)	3 (1%)				7 (2%)						
n.body	4 (1%)		5 (1%)										
n.possession							3 (1%)		4 (1%)				
V-to-N	**n.shape**	**n.artifact**	**n.state**	**n.communication**	**n.object**	**n.attribute**	**n.food**	**n.cognition**	**n.location**	**n.plant**	**n.act**	**n.body**	**n.group**
v.contact	13 (10%)	6 (4%)	9 (7%)		7 (5%)	4 (3%)	3 (2%)	2 (1%)	2 (1%)			2 (1%)	
v.communication		2 (1%)		6 (4%)									
v.change	5 (4%)		2 (1%)		4 (3%)	5 (4%)	3 (2%)			2 (1%)			
v.creation		9 (7%)		4 (3%)			2 (1%)	2 (1%)					
v.perception			3 (2%)										2 (1%)
v.motion	5 (4%)			2 (1%)							2 (1%)		
v.stative								2 (1%)					

Tab. 17: The combination of semantic primes to express the relation Undergoer. The grey cells show that for the N-to-V direction two noun primes (n.communication and n.artifact) account for most of the data, while for the V-to-N direction there is one noun prime (n.artifact) occurring with almost all verb primes in the data.

N-to-V	v.communication	v.contact	v.possession	v.creation	v.change	v.competition	v.motion	v.cognition	v.consumption	v.perception	v.social		
n.communication	23 (10%)	3 (1%)	3 (1%)	11 (5%)	4 (2%)			5 (2%)		3 (1%)	3 (1%)		
n.artifact		19 (8%)	3 (1%)	11 (5%)		4 (2%)	6 (3%)			3 (1%)			
n.possession			18 (8%)			3 (1%)							
n.plant		5 (2%)			8 (4%)								
n.animal		8 (4%)			3 (1%)	7 (3%)							
n.food		8 (4%)			4 (2%)	3 (1%)			3 (1%)				
n.substance		7 (3%)	4 (2%)		6 (3%)								
n.group		7 (3%)											
n.person		6 (3%)											
n.quantity		5 (2%)			3 (1%)								
n.attribute	5 (2%)				3 (1%)								
n.body		4 (2%)											
V-to-N	**n.communication**	**n.artifact**	**n.possession**	**n.food**	**n.body**	**n.substance**	**n.object**	**n.quantity**	**n.group**	**n.person**	**n.cognition**	**n.attribute**	**n.event**
v.possession		2 (2%)	9 (9%)										
v.contact		8 (8%)		2 (2%)				3 (3%)					
v.communication	4 (4%)									2 (2%)	2 (2%)		
v.consumption				(4%)									
v.body		2 (2%)			4 (4%)	3 (3%)							
v.motion		2 (2%)					3 (3%)						
v.competition		2 (2%)											2 (2%)
v.cognition	2 (2%)	2 (2%)									2 (2%)		
v.stative												2 (2%)	
v.social									2 (2%)				

Tab. 18: The combination of semantic primes to express the relation By-means-of. The grey cells show that only two noun primes (n.artifact and n.communication) are more frequent in the prime combinations, irrespective of the zero derivation direction, whereas the number of verb primes is much higher.

N-to-V	v.contact	v.communication	v.motion	v.creation	v.change	v.social	v.stative	v.perception	v.cognition	v.body	v.possession
n.artifact	67 (19%)	11 (3%)	14 (4%)	12 (3%)	11 (3%)		6 (2%)	4 (1%)		3 (1%)	
n.communication	3 (1%)	47 (13%)	3 (1%)	5 (1%)		7 (2%)			3 (1%)		3 (1%)
n.body	11 (3%)		4 (1%)								
n.substance	9 (3%)				4 (1%)						
n.cognition		4 (1%)				5 (1%)		4 (1%)	4 (1%)		
n.attribute		4 (1%)				3 (1%)					
n.relation	3 (1%)										
n.state						3 (1%)					

V-to-N	n.artifact	n.attribute	n.communication	n.possession	n.cognition	n.phenomenon	n.quantity	n.object	n.substance	n.state	n.shape
v.contact	19 (16%)	3 (3%)				3 (3%)		2 (2%)			2 (2%)
v.social		6 (5%)			2 (2%)						
v.communication		4 (3%)	5 (4%)							2 (2%)	
v.possession			3 (3%)	4 (3%)							
v.change	2 (2%)	4 (3%)					2 (2%)		2 (2%)		
v.emotion					4 (3%)						
v.creation	3 (3%)										
v.perception	2 (2%)	3 (3%)									
v.motion	2 (2%)					2 (2%)		2 (2%)			
v.stative	2 (2%)										

Tab. 19: The combination of semantic primes to express the relation Instrument. We notice that the noun prime n.artifact is basis for all verb primes and is the only prime to which zero nominalisations belong.

N-to-V	v.contact	v.creation	v.motion	v.body	v.change
n.artifact	131 (63%)	12 (6%)	9 (4%)	9 (4%)	9 (4%)
V-to-N	**n.artifact**				
v.contact	131 (75%)				

Tab. 20: The combination of semantic primes to express the relation Location. We show by means of the grey cells that two noun primes (n.artifact and n.location) prevail in combining with a higher number of verb primes (for N-to-V) or with the same number of verb primes (for V-to-N).

N-to-V	v.contact	v.emotion	v.stative	v.possession	v.social	v.creation
n.artifact	17 (14%)	11 (9%)	7 (6%)	9 (7%)	5 (4%)	5 (4%)
n.location	8 (6%)	13 (10%)	9 (7%)		6 (5%)	
n.object		10 (8%)				
V-to-N	**n.location**	**n.attribute**	**n.artifact**			
v.contact	7 (25%)					
v.change		3 (12%)				
v.motion	3 (12%)		3 (12%)			

Tab. 21: The combination of semantic primes to express the relation State. The grey cells show one noun prime (n.state) is more frequently combined with various verb primes, for each direction of the zero derivation.

N-to-V	v.contact	v.communication	v.emotion	v.change	v.stative	v.social
n.state	4 (13%)		3 (10%)	3 (10%)	2 (6%)	2 (6%)
n.quantity		3 (10%)				
n.feeling			3 (10%)			
n.attribute		2 (6%)				
V-to-N	**n.state**	**n.attribute**	**n.feeling**			
v.emotion	7 (29%)		2 (8%)			
v.change	3 (12.5%)	3 (12.5%)				
v.body	2 (8%)					

Tab. 22: The combination of semantic primes to express the relation Uses. The grey cells show, for N-to-V direction, the higher number of two noun primes (n.artifact and n.communication) and of two verb primes (v.contact and v.possession), while for the V-to-N direction one verb prime (v.contact) and one noun prime (n.artifact) prevail.

N-to-V	v.contact	v.possession	v.communication	v.body	v.change	v.creation	v.cognition	v.competition	v.stative
n.artifact	49 (40%)	36 (29.5%)		20 (16%)	6 (5%)	10 (8%)		6 (5%)	5 (4%)
n.substance	42 (34%)			11 (9%)					
n.communication	9 (7%)	5 (4%)	26 (21%)						
n.possession		9 (7%)							
n.time							8 (7%)		
n.attribute					6 (5%)				
n.food	6 (5%)				5 (4%)				
n.plant	5 (4%)								
V-to-N	**n.artifact**	**n.communication**	**n.substance**	**n.shape**	**n.attribute**	**n.possession**			
v.contact	13 (17%)	5 (7%)	4 (5%)	3 (4%)	2 (3%)				
v.body	7 (9%)								
v.creation	5 (7%)								
v.possession					2 (3%)	2 (3%)			
v.weather					2 (3%)				
v.change	2 (3%)								